

**Lokale asymptotische Theorie differenzierbar
parametrisierter statistischer Experimente,
insbesondere Markoff-Prozesse**

Dissertation

zur Erlangung des Grades eines Doktors der
Naturwissenschaften der Fakultät für Mathematik
der Ruhr-Universität Bochum

vorgelegt
von

Bernd Kuhle

1988



**Lokale asymptotische Theorie differenzierbar
parametrisierter statistischer Experimente,
insbesondere Markoff-Prozesse**

Dissertation

zur Erlangung des Grades eines Doktors der
Naturwissenschaften der Fakultät für Mathematik
der Ruhr-Universität Bochum

vorgelegt

von

Bernd Kuhle

1988

Danksagung

Die vorliegende Arbeit entstand in den Jahren 1987 und 1988 auf Anregung und unter Anleitung von Prof. Dr. V. Baumann am Institut für Mathematik an der Ruhr-Universität Bochum. Durch fruchtbare Diskussionen und wertvolle Ratschläge hat er die Fertigstellung dieser Arbeit wesentlich gefördert. Dafür gilt ihm mein ganz besonderer Dank.



Digitized by the Internet Archive
in 2026

Inhalt

1. Einführung
2. Lokale asymptotische Normalität von Markoff-Prozessen
3. Lokal schwache Konvergenz - differenzierbare Mannigfaltigkeiten als Parameterräume
- 3.I Lokal schwache Konvergenz
- 3.II Die Hauptsätze der Test- und Schätztheorie für lokal schwach konvergente Experimentenfolgen
- 3.III Lokale asymptotische Testtheorie
- 3.IV Lokale asymptotische Schätztheorie
4. Lokal schwache Konvergenz im Markoff-Fall
5. Anwendungen - Mannigfaltigkeiten als Parameterräume im Markoff-Fall
6. Differenzierbarkeit
7. Stochastische Entwicklung
8. Einige Aussagen über Markoff-Prozesse
9. Lebesgue-Zerlegung für Kerne
10. Differenzierbare Mannigfaltigkeiten
11. Beweistechnische Hilfsmittel

Notation

Sei P ein Mass auf einem messbaren Raum $(X, \mathcal{A})$, p ein Kern von $(X, \mathcal{A})$ nach einem messbaren Raum $(Y, \mathcal{B})$. Sofern keine Missverständnisse zu befürchten sind, schreiben wir

$$P(f) := \int P(dx) f(x)$$

für eine integrierbare Funktion f auf $(X, \mathcal{A})$ und

$$p(x, g(x, \cdot)) := \int p(x, dy) g(x, y)$$

für eine Funktion g auf $X \times Y$, die bei festem x integrierbar bzgl. des Masses $p(x, \cdot)$ ist.

Mit $F \circ P$ bezeichnen wir das Bildmass von P unter der messbaren Abbildung F .

Das Produkt zweier Masse P und Q notieren wir mit $P \otimes Q$. Entsprechend ist durch

$$P \otimes p(A \times B) := \int P(dx) 1_A(x) p(x, B),$$

für $A \times B \in X \times Y$ das Produkt des Masses P mit dem Kern p definiert.

Ist $(P_n : n \in \mathbb{N})$ eine Folge von Verteilungen und sind f_n , $n \in \mathbb{N}$ und f Zufallsvariable auf messbaren Räumen $(\Omega_n, \mathcal{A}_n)$ bzw. $(\Omega, \mathcal{A})$, so bedeutet die Schreibweise

$$f_n \longrightarrow f \quad (P_n),$$

dass f_n P_n -stochastisch gegen f konvergiert. Wir schreiben

$$f_n = o_{P_n}(a_n)$$

mit einer Folge reeller Zahlen a_n , falls

$$a_n^{-1} f_n \longrightarrow 0 \quad (P_n).$$

Für eine k -mal differenzierbare Abbildung f notieren wir die k -te Ableitung von f mit $D^k f$ oder auch $D_x^k f(x, \cdot)$, die k -te Ableitung an der Stelle y mit $D_y^k f$ oder $D_x^k f(x, \cdot) \Big|_{x=y}$.

1 Einführung

1.1 Asymptotisch normale Experimentenfolgen

Betrachte eine Folge messbarer Räume $(\Omega_n, \mathcal{A}_n)$, eine beliebige Menge T und eine Folge von Teilmengen $T_n \subseteq T$, mit der Eigenschaft $T_n \uparrow T$. Für $t \in T_n$ bezeichne $P_{n,t}$ eine Verteilung auf $(\Omega_n, \mathcal{A}_n)$. Die asymptotische Statistik untersucht die Folge der statistischen Experimente $E_n := (\Omega_n, \mathcal{A}_n, \{P_{n,t} : t \in T_n\})$. Ist D ein topologischer Raum, der Entscheidungsraum, versehen mit der Baire- σ -Algebra $\mathcal{B}_0(D)$, so ist eine Verlustfunktion eine Familie $W := (W_t : t \in T)$ von messbaren, nach unten beschränkten Funktionen $W_t : D \rightarrow \mathbb{R}$. Das Tripel (T, D, W) heisst Entscheidungsproblem. Eine randomisierte Entscheidungsfunktion (für das Experiment E_n und das Entscheidungsproblem (T, D, W)) ist ein Markoff-Kern ρ_n von $(\Omega_n, \mathcal{A}_n)$ nach $(D, \mathcal{B}_0(D))$, und die Grösse

$$R(\rho_n, t) := \int \int W_t(x) \rho_n(\omega_n, dx) P_{n,t}(d\omega_n)$$

heisst Risiko von ρ_n und t (bzgl. W). Die asymptotische Statistik hat das asymptotische Verhalten der Risikofunktion bei wachsendem $n \in \mathbb{N}$ zum Gegenstand. Für Untersuchungen in diesem Zusammenhang ist der Begriff der schwachen Konvergenz einer Folge statistischer Experimente grundlegend. Dieser Terminus wurde von Le Cam eingeführt. Verallgemeinert man die Definition der Entscheidungsfunktion in geeigneter Weise (vgl. Strasser (1)), so bedeutet die schwache Konvergenz der Folge der Experimente E_n gegen ein Grenzexperiment $E := (\Omega, \mathcal{A}, \{P_t : t \in T\})$ grob gesprochen folgendes: Für jede Folge verallgemeinerter Entscheidungsfunktionen β_n für E_n und (T, D, W) kann man die Risiken $R(\beta_n, t)$ für wachsende $n \in \mathbb{N}$ beliebig gut durch das Risiko $R(\beta, t)$ einer verallgemeinerten Entscheidungsfunktion β für das Grenzexperiment E und (T, D, W) approximieren, und zwar für jede stetige beschränkte Verlustfunktion W und jede endliche Teilmenge $\alpha \subseteq T$ (vgl. Strasser (1)). Sofern eine Folge statistischer Experimente schwach konvergiert, kann man also hoffen, Aussagen betreffs des asymptotischen Verhaltens der Folge der Risiken über entsprechende Aussagen für das Grenzexperiment zu gewinnen. Die sogenannten Hauptsätze der asymptotischen Test- und Schätztheorie,

die auf Le Cam zurückgehen (vgl. Strasser (1), 62, Strasser (2), 10), bestätigen diese Hoffnung.

Die statistische Theorie der Gauss-Shift-Experimente ist besonders weit entwickelt. Von daher ist klar, dass schwache Konvergenz gegen einen Gauss-Shift einen besonders glücklichen Fall darstellt. Was unter einem Gauss-Shift zu verstehen ist, präzisieren wir wie folgt.

1 Definition

$(\Omega, \mathcal{A}, P)$ sei ein Wahrscheinlichkeitsraum, $(H, \langle \cdot, \cdot \rangle)$ ein separabler Hilbert-Raum. Mit $L^2(P)$ notieren wir den Raum der quadratisch P -integrierbaren Funktionen auf Ω . Eine lineare Abbildung

$$L : H \longrightarrow L^2(P)$$

ist ein Standard-Gauss-Prozess für $(H, \langle \cdot, \cdot \rangle)$, falls

$$L(h) \circ P = N(0, |h|^2)$$

mit $|h|^2 := \langle h, h \rangle$. Ein statistisches Experiment $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ ist ein Gauss-Shift auf $(H, \langle \cdot, \cdot \rangle)$, wenn für die Dichtequotienten von P_h bzgl. P_0 gilt:

$$\frac{dP_h}{dP_0}(\omega) = \exp \left[L(h)(\omega) - \frac{1}{2} |h|^2 \right] P_0\text{-f.s.}$$

mit einem Standard-Gauss-Prozess L für $(H, \langle \cdot, \cdot \rangle)$ bzgl. P_0 . In diesem Fall nennen wir L einen zentralen Prozess für E (vgl. Strasser (1)). Ist speziell $(H, \langle \cdot, \cdot \rangle)$ ein euklidischer Raum, so existiert eine Zufallsvariable $X : \Omega \longrightarrow H$ derart, dass der zentrale Prozess eine Darstellung $L(h) = \langle h, X \rangle$ besitzt. Dann sprechen wir von X als der zentralen Zufallsvariablen des Gauss-Shifts E . Ihre Verteilungen genügen der Beziehung

$$X \circ P_h = N_H \circ \varepsilon_h$$

für die Faltung des Standard-Gauss-Masses N_H auf $(H, \langle \cdot, \cdot \rangle)$ mit der Einheitsmasse ε_h in $h \in H$ (vgl. Strasser (1), 5.2, 69.11). ■

Wir sind nun in der Lage, eine für unsere Zwecke geeignete Formulierung der schwachen Konvergenz gegen einen Gauss-Shift anzugeben.

2 Definition

Gegeben seien Teilmengen H_n des separablen Hilbert-Raumes H , die alle 0 enthalten und monoton von unten gegen H konvergieren. Die Folge der Experimente F_n soll durch

$$F_n := (\Omega_n, \mathcal{A}_n, \langle P_{n,t} : t \in H_n \rangle)$$

und der Gauss-Shift F vermöge

$$F := (\Omega, \mathcal{A}, \langle P_h : h \in H \rangle)$$

gegeben sein. Genau dann konvergiert die Folge $(F_n : n \in \mathbb{N})$ schwach gegen den Gauss-Shift F , wenn

$$\frac{dP_{n,h}}{dP_{n,0}}(\omega_n) = \exp \left[L_n(h)(\omega_n) - \frac{1}{2} \|h\|^2 + o_{P_{n,0}}(n^0) \right]$$

für eine Folge stochastischer Prozesse $(L_n(h) : h \in H)$ mit den Eigenschaften

$$L_n(h) * P_{n,0} \longrightarrow N(0, \|h\|^2) \text{ schwach für alle } h \in H \text{ und}$$

$$\alpha L_n(h_1) + \beta L_n(h_2) - L_n(\alpha h_1 + \beta h_2) = o_{P_{n,0}}(n^0) \quad \text{für } \alpha, \beta \in \mathbb{R} \text{ und } h_1, h_2 \in H.$$

Vereinfachend sagen wir auch, die Folge der Experimente ist asymptotisch normal. Die Folge der $(L_n(h) : h \in H)$ nennen wir eine zentrale Folge für $(F_n : n \in \mathbb{N})$. ■

Die Gleichwertigkeit dieser Definition mit anderen in der Literatur anzutreffenden ergibt sich mit Hilfe der Aussagen in Strasser (1), 80.2 und 80.9. Wenn schwache Konvergenz gegen einen Gauss-Shift vorliegt mit einer zentralen Folge $(L_n(h) : h \in H)$, so ist eine weitere Folge linearer Prozesse $(G_n(h) : h \in H)$ dann und nur dann zentral, wenn

$$L_n(h) - G_n(h) = o_{P_{n,0}}(n^0) \text{ für alle } h \in H$$

gilt. Die zentrale Folge $(L_n(h) : h \in H)$ kann stets als Folge linearer Prozesse gewählt werden. Insbesondere existiert eine Folge von Zufallsvariablen $X_n : \Omega_n \longrightarrow H$ mit $L_n(h) = \langle X_n, h \rangle$, sofern H ein euklidischer Raum ist.



1.II Die lokale asymptotische Betrachtungsweise

Aus dem Bestreben, Folgen statistischer Entscheidungsfunktionen für ein Entscheidungsproblem hinsichtlich ihrer Qualität zu vergleichen, ist die lokale asymptotische Betrachtungsweise in der Statistik hervorgegangen. Wir gehen aus von einer Folge von Experimenten $F_n := (\Omega_n, \mathcal{A}_n, (P_{n,t} : t \in H_0))$, wobei H_0 eine offene Teilmenge eines separablen Hilbert-Raums H sei. Betrachte ein Testproblem (H, K) für H_0 und definiere die Menge J als den topologischen Rand von H in H_0 . Eine Folge kritischer Funktionen $\varphi_n : \Omega_n \rightarrow [0,1]$ heisst asymptotisch α -ähnlich auf J , falls $\lim_{n \rightarrow \infty} P_{n,h} \varphi_n = \alpha$ für alle $h \in J$. Sollen nun zwei solche Testfolgen $\varphi_{1,n}$ und $\varphi_{2,n}$ hinsichtlich ihrer Qualität gegenübergestellt werden, so ist es naheliegend, die Grenzwerte $\lim_{n \rightarrow \infty} P_{n,h} \varphi_{i,n}$, $i = 1, 2$ zu vergleichen. Nun sind aber bekanntermassen viele Testfolgen konsistent, erfüllen also

$$\lim_{n \rightarrow \infty} P_{n,h} \varphi_n = 0, \text{ wenn } h \in H \setminus J \text{ und}$$

$$\lim_{n \rightarrow \infty} P_{n,h} \varphi_n = 1, \text{ falls } h \in K \setminus J.$$

Konsistente, α -ähnliche Testfolgen lassen sich also über das asymptotische Verhalten ihrer Gütefunktionen nicht unterscheiden. Eine Alternative besteht darin, die Grenzwerte $\lim_{n \rightarrow \infty} P_{n,h_n} \varphi_n$ für Folgen $(h_n : n \in \mathbb{N}) \subset H_0$ zu untersuchen, insbesondere wenn die h_n gegen ein Element $h_0 \in J$ streben. Denn das Verhalten der Gütefunktionen für solche Folgen gibt Auskunft darüber, wie gut die Testfolge naheliegende Alternativen trennt. Ist nun die Folge der Experimente hinreichend regulär, (nämlich in h_0 lokal schwach konvergent gegen einen Gauss-Shift, vgl. 3.7,) so gilt für Folgen mit $n^{1/2}(h_n - h_0) \rightarrow 0$ und beliebige Testfolgen $(\varphi_n : n \in \mathbb{N})$:

$$\lim_{n \rightarrow \infty} \left[P_{n,h_n} \varphi_n - P_{n,h_0} \varphi_n \right] = 0.$$

Solche Folgen können also ebenfalls nicht für asymptotische Qualitätsbetrachtungen dienen. Anders ist die Situation, wenn $n^{1/2}(h_n - h_0)$ konvergiert und der Grenzwert von 0 verschieden ist. Tatsächlich existieren dann in hinreichend gutartigen Fällen, nämlich wenn die Folge der Experimente in h_0 lokal asymptotisch normal ist (, vgl. 1.III und 3.III), Unterschiede im asymptotischen Verhalten von $P_{n,h_n} \varphi_n$ für verschiedene Folgen kritischer Funktionen φ_n , und man kann Schranken für die

Haufungspunkte dieser Folgen angeben.

Im Rahmen der asymptotischen Schätztheorie lässt sich der lokale Standpunkt ebenfalls motivieren.

Wir betrachten eine offene Teilmenge H_0 eines euklidischen Raums H und eine Folge statistischer Experimente

$$F_n := (\Omega_n, \mathcal{A}_n, (P_{n,h} : h \in H_0)).$$

Eine Folge von Schätzern $\hat{h}_n$ für den Parameter h nennen wir konsistent, falls

$$\lim \hat{h}_n = h \quad (P_{n,h}).$$

Für jede stetige beschränkte Verlustfunktion W gilt deshalb

$$\lim P_{n,h} [W_h(\hat{h}_n)] = W(h),$$

und dieser Grenzwert ist zur Unterscheidung zweier konsistenter Schätzfolgen also ungeeignet. Eine schärfere Forderung an die Schätzfolge ist die der schnellen Konsistenz. Wir nennen die Folge der $\hat{h}_n$ schnell konsistent, wenn zu jedem $h \in H_0$ und jedem $\varepsilon > 0$ eine Konstante $C(\varepsilon)$ existiert, so dass

$$\limsup_{n \in \mathbb{N}} P_{n,h} \left\{ n^{1/2} |\hat{h}_n - h| > C(\varepsilon) \right\} < \varepsilon.$$

Schnelle Konsistenz ist in sofern eine hervorragende Eigenschaft, als im allgemeinen höhere Konsistenzgeschwindigkeiten nicht erwartet werden können. Genauer gilt: Ist die Folge der Experimente F_n in h_0 lokal asymptotisch normal, so gibt es keine Schätzfolge, für die $n^{1/2} |\hat{h}_n - h|$ in h_0 lokal gleichmässig $P_{n,h}$ -stochastisch gegen 0 konvergiert (vgl. 3.7).

Betrachte nun das asymptotische Verhalten der Folge $n^{1/2} |\hat{h}_n - h|$. Fisher schlug vor, die Konkurrenz um die effiziente Schätzfolge auf den Kreis der asymptotisch zentral normalverteilten Schätzfolgen zu beschränken und jene effizient zu nennen, deren Grenzverteilung eine minimale Varianz besitzt. Die Unzulänglichkeiten des Fisherschen Effizienzbegriffs wurden durch die Beispiele supereffizienter Schätzer von Hodges (vgl. Ibragimov/Has'minskii) und die Kritik von Wolfowitz (vgl. Wolfowitz) verdeutlicht. Eine gleichmässige untere Schranke für die Varianzen der Grenzverteilungen asymptotisch normaler Schätzfolgen existiert nicht.

In der lokalen asymptotischen Theorie fixiert man ein beliebiges

$h_0 \in H_0$. Die Grösse $n^{1/2}|\hat{h}_n - h_0|$ ist dann ein Schätzer für $n^{1/2}|h - h_0| =: t$. Natürlich können nur solche t geschätzt werden, für die t beschränkt bleibt. Die lokale asymptotische Theorie vergleicht nun Schätzfolgen dadurch, dass sie feststellt, wie gut $n^{1/2}|\hat{h}_n - h_0|$ den Parameter t schätzt. Für eine Verlustfunktion W von der Form $W_h(x) := W(x - h)$ betrachten wir also das Verhalten von

$$(1) \quad P_{n, h_n} \left[W \left(n^{1/2}(\hat{h}_n - h_0) - n^{1/2}(h - h_0) \right) \right] =$$

$$P_{n, h_n} \left[W \left(n^{1/2}(\hat{h}_n - h) \right) \right]$$

für $h_n \in H$, so dass $n^{1/2}(h_n - h_0)$ konvergiert für $n \rightarrow \infty$.

Eine Schätzfolge ist dann lokal asymptotisch optimal in h_0 , wenn sie die Grenzwerte unter (1) in geeigneter Weise optimiert. Natürlich wird man vor allem an Aussagen bei beliebiger Wahl des Punktes h_0 interessiert sein, denn der wahre Parameter h ist dem Statistiker ja nicht bekannt. Die lokale Betrachtungsweise entspricht insofern der Praxis, als der wahre Parameter h ohnehin mit grosser Wahrscheinlichkeit lokalisiert werden kann. Mit wachsendem Stichprobenumfang sind deshalb gerade die kleinen Abweichungen interessant. Genau diese werden durch die lokale Sichtweise untersucht.

Tatsächlich liefert die lokale asymptotische Betrachtungsweise Aussagen über das asymptotische Verhalten der Grössen (1), wenn die Folge der Experimente in h_0 hinreichend regulär ist, genauer lokal asymptotisch normal ist.

1.III Lokale asymptotische Normalität

Die Überlegungen zur lokalen Betrachtungsweise in der asymptotischen Statistik haben uns veranlasst, das asymptotische Verhalten der Risiken von Folgen statistischer Entscheidungsfunktionen unter Verteilungen P_{n,h_n} zu analysieren, wenn $n^{1/2}(h_n - h_0)$ konvergiert. Für eine offene Teilmenge H_0 eines separablen Hilbert-Raumes H sei

$$(1) \quad F_n := (\Omega_n, \mathcal{F}_n, \langle P_{n,h} : h \in H_0 \rangle)$$

eine Folge statistischer Experimente. Wähle irgendein $h_0 \in H_0$ aus und definiere Mengen H_n durch

$$(2) \quad H_n := \left\{ h \in H : h_0 + n^{-1/2}h \in H_0 \right\}$$

sowie Experimente durch

$$(3) \quad F_{n,h_0} := \left[\Omega_n, \mathcal{F}_n, \left\{ P_{n,h_0 + n^{-1/2}h} : h \in H_n \right\} \right].$$

3 Definition

Die Folge der Experimente (1) heisst in h_0 lokal asymptotisch normal, wenn die Folge der Experimente (3) schwach gegen einen Gauss-Shift auf $(H, \langle \cdot, \cdot \rangle_{h_0})$ konvergiert. Das Skalarprodukt darf dabei von h_0 abhängen. Die Experimente (3) nennen wir die in h_0 lokalisierten Experimente. ■

Wie in 1.I und 1.II dargestellt, kann man das Problem der lokalen asymptotischen Schranken und der lokal asymptotisch optimalen Entscheidungsfunktionen im Fall von lokaler asymptotischer Normalität auf die entsprechenden Fragestellungen für das Grenzexperiment zurückführen. Für den Fall unabhängiger, identischer Versuchswiederholungen, wenn also die Folge der Experimente durch

$$F_n := (\Omega^n, \mathcal{F}^n, \langle P_h^n : h \in H_0 \rangle)$$

gegeben ist, sind diesbezügliche Arbeiten von verschiedenen Autoren erschienen, basierend auf Ideen von Le Cam und Hajek. Darstellungen hierzu findet man beispielsweise in Strasser (1), Strasser (2) und Ibragimov/Has'minskii. Als ein hinreichendes und im wesentlichen auch notwendiges Kriterium für die lokale

asymptotische Normalität hat sich dabei eine Regularitätsbedingung an die Verteilungsklasse $(P_h : h \in H_0)$ erwiesen, nämlich deren L^2 -Differenzierbarkeit.

Weitere Arbeiten zu diesem Thema stammen von Pfanzagl (vgl. Pfanzagl (1) und (2)). Der Autor geht von anderen Differenzierbarkeitsbegriffen für Kurven aus. Die von Pfanzagl eingeführte DCC-Differenzierbarkeit ist ein Name für einen Satz von Bedingungen, der dem DCC-Kriterium aus Loève entnommen ist und sehr nahe an den beweistechnisch notwendigen Eigenschaften für stochastische Entwicklungen von Dichtequotienten liegt. Der DCC-Differenzierbarkeitsbegriff ist weitgehend gleichwertig mit der L^2 -Differenzierbarkeit (vgl. Pfanzagl (2)) von Kurven von Verteilungen. Pfanzagls Begriffsbildung erlaubt Verallgemeinerungen auf stärkere Differenzierbarkeitseigenschaften, die stochastische Entwicklungen höherer Ordnung ermöglichen.

1.IV Uebersicht

In dieser Arbeit betrachten wir Folgen statistischer Experimente, die durch einen stationären oder zeitlich homogenen Markoff-Prozess beschrieben werden. Genauer sei $(E, \mathcal{B}(E))$ ein polnischer Raum, $H_0 \subseteq \mathbb{H}$ eine offene Teilmenge eines separablen Hilbert-Raums, und für jedes $h \in H_0$ bezeichne p_h einen Markoff-Kern auf $(E, \mathcal{B}(E))$ sowie $P_{1,h}$ eine Verteilung auf $(E, \mathcal{B}(E))$. Wir definieren $(E^n, \mathcal{B}(E^n))$ als das n -fache Produkt des messbaren Raumes $(E, \mathcal{B}(E))$ mit sich selbst und Verteilungen $P_{n,h}$ als die eindeutig bestimmten Fortsetzungen der Mengenfunktionen

$$P_{n,h}(B_1 \times \dots \times B_n) :=$$

$$\int P_{1,h}(dx_1) 1_{B_1}(x_1) \int p_h(x_1, dx_2) 1_{B_2}(x_2) \dots \int p_h(x_{n-1}, dx_n) 1_{B_n}(x_n),$$

$B_1, \dots, B_n \in \mathcal{B}(E)$, zu Wahrscheinlichkeitsmassen auf $(E^n, \mathcal{B}(E^n))$.

Wir untersuchen die Folge der Experimente

$$(1) \quad F_n := (E^n, \mathcal{B}(E^n), \{P_{n,h} : h \in H_0\}).$$

Der 6. Abschnitt enthält die Definitionen und Eigenschaften L^2 -differenzierbarer Kurven und Familien von Markoff-Kernen. Vorausgesetzt wird hier, wie in der ganzen Arbeit, dass die beteiligten Markoff-Kerne Dichten bzgl. eines weiteren Kernes besitzen. Die von uns geprägten Begriffe der L^2 -Differenzierbarkeit stellen Verallgemeinerungen der entsprechenden Begriffe für Kurven und Familien von Verteilungen dar (vgl. Strasser (1), Witting, und 6.20), wie sie von Le Cam geprägt wurden. Ebenso ist eine Verallgemeinerung der DCC-Differenzierbarkeit auf den Markoff-Fall möglich. Der in dieser Arbeit gewählte Weg stellt keine Bewertung der beiden Begriffsbildungen dar, sondern rührt her von der Historie meines persönlichen Bekanntwerdens mit dem Problemkreis, dem diese Arbeit angehört.

Die L^2 -Differenzierbarkeit einer Kurve $s \rightarrow p_s$ von Markoff-Kernen kann als die wesentliche Voraussetzung dienen, unter der ein stochastische Entwicklung der Dichtequotienten der $n-1$ -Schritt Übergangskerne $p_{n-1,s}$ bzgl. $p_{n-1,0}$ zulässig ist und zwar

$$(2) \quad \frac{dp_{n-1, n-1/2}^{u_n}}{dp_{n-1, 0}}(x_1, \dots, x_n) =$$

$$\exp \left\{ n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}[\xi^2] + o_{P_{n,0}}(n^0) \right\}$$

für konvergente Folgen u_n . $P_{2,0}$ ist dabei aus einer stationären Verteilung $P_{1,0}$ von p_0 und p_0 selbst gebildet (vgl. 2.1). In wichtigen Spezialfällen ist die L^2 -Differenzierbarkeit alleine schon hinreichend für die Gültigkeit von (2) (vgl. 2.2). Der Beweis dieser Aussage benutzt einen allgemeinen Satz zur stochastischen Entwicklung von Dichtequotienten, der im 7. Abschnitt formuliert und bewiesen wird. Er betrifft die Randverteilungen beliebiger, nicht notwendig zeitlich homogener Markoff-Prozesse und scheint uns von gewissem eigenständigen Interesse zu sein. Die Aussage (2) verwenden wir in den Sätzen 2.3 bis 2.7, um hinreichende Kriterien für die lokale asymptotische Normalität der Experimentenfolge (1) zu gewinnen.

Der Rest des 2. Abschnitts ist der Frage gewidmet, ob aus der lokalen asymptotischen Normalität einer Folge von Experimenten von Typ (1) auf die L^2 -Differenzierbarkeit von Kurven geschlossen werden kann. Ein Hauptresultat in diesem Zusammenhang ist, dass eine Beziehung wie (2) für eine Kurve $s \rightarrow p_s$ genau dann besteht, wenn die endlich-dimensionalen Randverteilungen für p_s und die stationäre Verteilung zu p_0 beliebig nahe an den entsprechenden Verteilungen einer differenzierbaren Kurve durch p_0 liegen, gemessen im Hellinger-Abstand (vgl. 2.11). Fordert man gewisse weitere Eigenschaften ausser der Gültigkeit von (2), so kann man sogar direkt auf die Differenzierbarkeit der Kurve schliessen. Dies legen wir in den letzten Aussagen des 2. Abschnitts dar.

Betrachte nun eine differenzierbare Mannigfaltigkeit M in einem Banach-Raum B (vgl. 10) und eine Familie von beliebigen Experimenten

$$(3) \quad F_n := (\Omega_n, \mathcal{A}_n, (P_{n,m} : m \in M)).$$

F_n ist hier also nicht notwendig vom Markoff-Typ. Das Konzept der

lokalen asymptotischen Normalität soll auf diesen Fall übertragen werden. Zu diesem Zweck ist zunächst ein geeigneter Begriff der Konvergenz der Experimentenfolge gegen einen Gauss-Shift anzugeben.

Differenzierbare Hilbert-Mannigfaltigkeiten, wie wir sie im 10. Abschnitt bestimmen, haben in jedem Punkt $m \in M$ Tangentialräume $T_m M$, die Teilmengen des Banach-Raums B sind und eine Hilbert-Raum-Struktur mit einem gewissen Skalarprodukt $\langle \cdot, \cdot \rangle_m$ tragen. Wir definieren im Unterabschnitt 3.I, was es bedeutet, wenn die Folge der Experimente F_n in einem $m \in M$ lokal schwach gegen den Gauss-Shift auf $(T_m M, \langle \cdot, \cdot \rangle_m)$ konvergiert. Die so definierte Konvergenzeigenschaft hängt nicht von dem Atlas - der Parametrisierung - der Mannigfaltigkeit ab. Sie ist eine Verallgemeinerung eines leicht verschärften Begriffs der schwachen Konvergenz einer Folge von lokalisierten Experimenten, wenn M eine offene Teilmenge eines Hilbert-Raumes ist (vgl. 2.5). Lokal schwache Konvergenz in $m \in M$ impliziert, dass zwei Folgen von Verteilungen $(P_{n,m} : n \in \mathbb{N})$, $i = 1, 2$, wechselseitig benachbart sind, wenn $n^{1/2}(m_{i,n} - m)$ konvergiert (vgl. 3.6). Die bereits erwähnten Aussagen 3.7 haben dann Gültigkeit.

Mit den so präzisierten Vorstellungen der Konvergenz von Experimentenfolgen vom Typ (3) sind wir in der Lage, lokale Versionen der sogenannten Hauptsätze der Test- und Schätztheorie für solche Experimentenfolgen zu formulieren und zu beweisen. Dies geschieht im Unterabschnitt 3.II. Sei die Folge der Experimente (3) in $m \in M$ lokal schwach konvergent gegen den Gauss-Shift auf $T_m M$ und $(m_n : n \in \mathbb{N}) \subset M$, so dass $n^{1/2}(m_n - m) \rightarrow t$. Ferner sei $N_1 \subseteq \mathbb{N}$ irgendeine Teilfolge. Einem Testproblem (H, K) für M entspricht dann ein Testproblem $(\hat{H}, \hat{K})$ für $T_m M$, und für eine beliebige Folge kritischer Funktionen φ_n für F_n und (H, K) existiert ein Test φ für F und $(\hat{H}, \hat{K})$ mit

$$\limsup_{n \in N_1} P_{n,m_n} \varphi_n \geq P_t \varphi, \quad t \in \hat{H} \text{ und}$$

$$\limsup_{n \in N_1} P_{n,m_n} \varphi_n \leq P_t \varphi, \quad t \in \hat{K}.$$

Entsprechend existiert für jede Schätzfolge einer

differenzierbaren Abbildung f auf M eine randomisierte Schätzfunktion ρ für Df_m , einer Abbildung auf $Tg_m M$, mit

$$\limsup_{n \in \mathbb{N}_1} P_{n,m_n} \left[W \left[n^{1/2} (\hat{f}_n - f(m_n)) \right] \right] \geq \iint W \left[x - Df_m(t) \right] \rho(\omega, dx) P_t(d\omega).$$

Die Resultate aus 3.II wenden wir in 3.III und 3.IV an, um für verschiedene Test- und Schätzprobleme lokal asymptotische Lösungen zu bestimmen.

Im 4. Abschnitt machen wir die Ergebnisse zur lokal schwachen Konvergenz für den Markoff-Fall verfügbar. Wir betrachten Folgen von Experimenten

$$(4) \quad F_n := (E^n, \mathcal{B}(E^n), (P_{n,m} : m \in M)).$$

Die Masse $P_{n,m}$ sind dabei wie für den 2. Abschnitt aus einem Übergangskern p_m und einer Startwahrscheinlichkeit $P_{1,m}$ gebildet. m variiert in einer differenzierbaren Hilbert-Mannigfaltigkeit M . Auch in diesem Fall erweist sich eine geeignet definierte L^2 -Differenzierbarkeit der Familie $\{p_m : m \in M\}$ (vgl. 6.22) als zentrale hinreichende Bedingung für die lokale schwache Konvergenz der Experimentenfolge F_n .

Anwendungen finden sich im 5. Abschnitt. Wir behandeln Familien $p(M)$ von Markoff-Kernen, die durch eine euklidische Mannigfaltigkeit M in einem Banach-Raum parametrisiert sind. Wir geben einen Satz von Bedingungen an, unter denen eine Experimentenfolge wie (4) überall lokal schwach gegen einen Gauss-Shift konvergiert. Diese Bedingungen sind weitgehend unabhängig von dem speziellen Atlas der Mannigfaltigkeit M . Die lokal zentralen Zufallsvariablen haben eine Darstellung $n^{1/2}(X_n - m)$, $m \in M$, $n \in \mathbb{N}$ mit von m unabhängigen Zufallsvariablen X_n . X_n ist im Prinzip ein Maximum-Likelihood-Schätzer. Für Experimente, die unsere Bedingungen erfüllen, leiten wir statistische Entscheidungsfunktionen mit guten globalen und lokalen Eigenschaften her. Dazu greifen wir auf bereits bekannte Resultate aus Billingsley (1) zurück. Ein grosser Teil unserer Ergebnisse ist mit entsprechenden Resultaten in Billingsley (1) verwandt.

Die restlichen Abschnitte 8. - 11. enthalten beweistechnische

Details, wie etwa Aussagen zur Lebesgue-Zerlegung von Markoff-Kernen, einige Lemmata über Markoff-Prozesse sowie die in dieser Arbeit zugrundeliegende Konzeption einer differenzierbaren Mannigfaltigkeit. Wir betonen, dass diese von der üblicherweise in der Literatur anzutreffenden abweicht. Sie ist allerdings ein Spezialfall des Mannigfaltigkeitsbegriffs, wie er beispielsweise in Choquet-Bruhart gegeben wird. Unser Begriff einer differenzierbaren Mannigfaltigkeit ist inspiriert von einer Konstruktion, wie sie in Spallek für den $\mathbb{R}^n$ durchgeführt wurde.

2. Lokale asymptotische Normalität von Markoff-Prozessen

Gegenstand dieser Arbeit sind Folgen statistischer Experimente, bei denen die Folge der Beobachtungen Realisierungen eines Markoff-Prozesses sind. Wir werden untersuchen, unter welchen Bedingungen eine solche Folge lokal asymptotisch normal ist. Genauer betrachten wir die folgende Situation:

H bezeichne eine Teilmenge eines Hilbert-Raumes $\mathbb{H}$, deren topologisches Innere nicht-leer ist. $(E, \mathcal{B}(E))$ sei ein polnischer Raum E versehen mit der σ -Algebra $\mathcal{B}(E)$ seiner Borel-Mengen, r sei ein Kern auf $(E, \mathcal{B}(E))$. Mit $p(r)$ notieren wir die Menge aller Markoff-Kerne auf $(E, \mathcal{B}(E))$, die eine Dichte bzgl. r besitzen (vgl. 9.1). Auf $(E, \mathcal{B}(E))$ möge eine Familie

$$p(H) := \left\{ p_h : h \in H \right\} \subseteq p(r)$$

von Markoff-Kernen gegeben sein. Dann existiert die Lebesgue-Zerlegung von p_h bzgl. $p_{h'}$ für beliebige $h, h' \in H$ (vgl. 9.5 - 9.8), die wir mit

$$\left\{ \frac{dp_h}{dp_{h'}}, \left\{ N_h(x) : x \in E \right\} \right\}$$

bezeichnen. Mit Hilfe der Markoff-Kerne p_h wird gemäss dem Satz von Ionescu Tulcea vermöge

$$p_{n-1,h} \left[x, B_2 x \dots x B_n \right] := \int_{B_2} p_h(x, dx_2) \int_{B_3} p_h(x_2, dx_3) \dots \int_{B_n} p_h(x_{n-1}, dx_n)$$

für $B_2, \dots, B_n \in \mathcal{B}(E)$ und $x \in E$ eindeutig ein Markoff-Kern von $(E, \mathcal{B}(E))$ nach $(E^{n-1}, \mathcal{B}(E^{n-1}))$ festgelegt. Ist des weiteren $\left\{ p_{1,h} : h \in H \right\}$ eine Familie von Verteilungen auf $(E, \mathcal{B}(E))$, so wird durch

$$\begin{aligned} P_{n,h} \left[B_1 x \dots x B_n \right] &:= \\ P_{1,h} \otimes p_{n-1,h} \left[B_1 x \dots x B_n \right] &:= \\ \int_{B_1} p_{n-1,h} \left[x_1, B_2 x \dots x B_n \right] P_{1,h}(dx_1) & \end{aligned}$$

eine Familie von Verteilungen $P_{n,h}$ auf $\langle E^n, \mathcal{B}(E^n) \rangle$ definiert. Für jedes $h \in H$ existiert der projektive Limes P_h für $\langle P_{n,h} : n \in \mathbb{N} \rangle$. Ist

$$\langle E^\infty, \mathcal{B}(E^\infty) \rangle := \bigotimes_{n=1}^{\infty} \langle E, \mathcal{B}(E) \rangle$$

und

$$X_i : E^\infty \longrightarrow E$$

die Projektion auf die i -te Komponente, so ist durch

$$\left[E^\infty, \mathcal{B}(E^\infty), P_h, \left[X_i : i \in \mathbb{N} \right] \right]$$

der durch $P_{1,h}$ und p_h bestimmte kanonische Markoff-Prozess gegeben, und es ist

$$P_{n,h} = \left[X_1, \dots, X_n \right] * P_h.$$

Wir werden die Folge der Experimente

$$F_n := \left[E^n, \mathcal{B}(E^n), \left\{ P_{n,h} : h \in H \right\} \right]$$

untersuchen. Ihr entspricht eine Folge von Stichprobenvariablen, die einen zeitlich homogenen Markoff-Prozess mit Übergangswahrscheinlichkeiten p_h und Startverteilung $P_{1,h}$ bilden.

Zur Abkürzung definieren wir noch für $1 \leq i_1 \leq \dots \leq i_k \leq n$, $n \in \mathbb{N}$,

$$P_{i_1, \dots, i_k, h} \left[B_1 x_1 \dots x_k B_k \right] :=$$

$$P_{n,h} \left[C_1 x_1 \dots x_k C_n \right]$$

mit

$$C_i := \begin{cases} B_j & , i = i_j \\ E & , \text{sonst} \end{cases}$$

und bezeichnen mit $P_{i_1, \dots, i_k, h}$ das durch obige Vorschrift eindeutig festgelegte Mass auf $\langle E^k, \mathcal{B}(E^k) \rangle$.

In den folgenden Sätzen werden wir hinreichende Bedingungen für die lokale asymptotische Normalität der Folge $\left[F_n : n \in \mathbb{N} \right]$ statistischer Experimente in einem Fusspunkt h' aus dem topologischen Inneren von H herleiten. Als wesentliche Eigenschaft der Familie wird sich dafür deren Differenzierbarkeit erweisen. In wichtigen Spezialfällen reicht sie alleine schon hin, um die lokale asymptotische Normalität der Folge sicher zu stellen.

Unter einer Kurve in $p(r)$ verstehen wir im folgenden stets eine Abbildung

$$(-c, c) \ni s \longrightarrow p_s \in p(r)$$

mit einem geeigneten, positiven $c \in \mathbb{R}$.

1 Satz

Gegeben sei eine Kurve $s \longrightarrow p_s$ in $p(r)$. $P_{1,0}$ bezeichne eine stationäre Verteilung von p_0 . Wenn die Kurve $L^2(P_{2,0})$ -differenzierbar mit Tangentialvektor ξ ist (vgl. 6.5 und 6.8) und

$$(1) \quad \frac{1}{n} \sum_{i=1}^n P_0 \left[x_i, \xi^2(x_i, \cdot) \right] \longrightarrow P_{2,0} \left[\xi^2 \right] \quad (P_{n,0})$$

erfüllt, so ist für jede Folge $(u_n : n \in \mathbb{N}) \subset \mathbb{R} \setminus \{0\}$ mit $u_n \longrightarrow u'$

$$(2) \quad \frac{dp_{n-1, u_n n^{-1/2}}}{dp_{n-1,0}} =$$

$$\exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0} \left[\xi^2 \right] + o_{P_{n,0}}(n^0) \right]$$

und

$$(3) \quad \left[n^{-1/2} \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) \right] * P_{n,0} \longrightarrow N(0, P_{2,0} \left[\xi^2 \right])$$

Beweis

Wir wollen für den Beweis dieser Aussage den Satz über die stochastische Entwicklung von Dichtequotienten (vgl. 7.4) heranziehen. Dazu ist nachzuweisen, dass

$$(4) \quad n P_{2,0} \left[\left\{ 2 \left[\frac{dp_{u_n n^{-1/2}}}{dp_0} \right]^{1/2} - 2 - n^{-1/2} u_n \xi \right\}^2 \right] \longrightarrow 0,$$

$$(5) \quad n P_{2,0} \left[\left\{ n^{-1/2} u_n \xi \right\}^2 \right] \leq c < \infty \text{ für alle } n \in \mathbb{N}, \text{ und ein } c,$$

$$(6) \quad n P_{2,0} \left[\left\{ n^{-1/2} u_n \xi \mathbf{1}_{\left\{ |n^{-1/2} u_n \xi| > \varepsilon \right\}} \right\}^2 \right] \longrightarrow 0$$

für alle $\varepsilon > 0$. Denn aus (5) und (6) folgt, dass

$$\left\{ n^{-1/2} u_n \xi(X_i, X_{i+1}) : 1 \leq i \leq n-1, n \geq 2 \right\}$$

ein Lindeberg-Feld für $\left\{ E^n, \mathcal{B}(E^n), P_{n,0} : n \in \mathbb{N} \right\}$ ist. Dabei hat

man nur $\left\{ P_{n,0} \right\}_{i,i+1} = P_{2,0}$ für alle $1 \leq i < n$ zu beachten. Die

Bedingung (4) entspricht in unserem (stationären) Fall gerade der Bedingung (1) aus 7.4. Offenbar gilt (4), wenn die Differenzierbarkeit der Kurve $s \rightarrow p_s$ mit Tangentialvektor g Berücksichtigung findet. Die Ungleichungen (5) sind erfüllt, weil $g \in L^2(P_{2,0})$ und $(u_n : n \in \mathbb{N})$ beschränkt ist. Aus denselben Gründen und weil

$$n^{1/2} u_n^{-1} \varepsilon \rightarrow \infty$$

gilt (6), sofern $\lim u_n \neq 0$. Im anderen Fall, $\lim u_n = 0$, ist die Konvergenz in (6) leicht einzusehen. Mithin haben wir

$$\begin{aligned} (7) \quad & \frac{dp_{n-1, u_n^{-1/2}}}{dp_{n-1,0}} = \\ & \exp \left\{ n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \right. \\ & \quad \left. n^{-1/2} u_n \sum_{i=1}^{n-1} p_0 \left[x_i, g(x_i, \cdot) \right] - \right. \\ & \quad \left. \frac{u_n^2}{2n} \sum_{i=1}^{n-1} p_0 \left[x_i, g^2(x_i, \cdot) \right] - \right. \\ & \quad \left. \sum_{i=1}^{n-1} p_{u_n^{-1/2}} \left\{ x_i, N_{u_n^{-1/2}}(x_i) \right\} + o_{P_{n,0}}(n^0) \right\}. \end{aligned}$$

Laut Satz 6.3 ist

$$p_0(x_i, g(x_i, \cdot)) = 0 \quad P_{1,0}\text{-f.s.}$$

Also verschwindet die zweite Summe im Argument der Exponentialfunktion in (7). Die dritte Summe dort strebt voraussetzungsgemäss gegen

$$\frac{u_n^2}{2} P_{2,0} \left[g^2 \right].$$

Aufgrund des Satzes 6.7 wissen wir schliesslich, dass die letzte Summe auf der rechten Seite von (7) $P_{n,0}$ -stochastisch gegen 0 konvergiert. Hieraus ergibt sich die Aussage (2) des Satzes.

Die schwache Konvergenz unter (3) legen wir unter Zuhilfenahme des Satzes von Brown (vgl. Gässler/Stute, 9.2.3, 9.2.8) dar.

$$(8) \quad \left[n^{-1/2} \varepsilon^2 \langle X_i, X_{i+1} \rangle : 1 \leq i \leq n-1; n \geq 2 \right]$$

ist ein Martingaldifferenzschema bzgl.

$$\left[\sigma \langle X_1, \dots, X_i \rangle : 1 \leq i \leq n-1; n \geq 2 \right]$$

in Folge der Markoff-Eigenschaft und wegen 6.3. Ferner gilt

$$(9) \quad \sum_{i=1}^{n-1} P_0 \left[n^{-1} \varepsilon^2 \langle X_i, X_{i+1} \rangle \mid X_1, \dots, X_1 \right] = \\ n^{-1} \sum_{i=1}^{n-1} P_0 \left[P_0 \left[\varepsilon^2 \langle X_i, \cdot \rangle \mid X_1, \dots, X_1 \right] \right] \xrightarrow{d} P_{2,0} \left[\varepsilon^2 \right] \quad (P_0)$$

gemäss Voraussetzung (1) und schliesslich ist

$$(10) \quad P_0 \left[\left| \sum_{i=1}^{n-1} P_0 \left\{ n^{-1} \varepsilon^2 \langle X_i, X_{i+1} \rangle \mid \left| n^{-1} \varepsilon^2 \langle X_i, X_{i+1} \rangle \right| > \varepsilon \right\} \right| \right. \\ \left. \left. \mid X_1, \dots, X_1 \right| \right] = \\ \frac{n-1}{n} P_{2,0} \left[\varepsilon^2 \mathbf{1}_{\{|\varepsilon| > \varepsilon n^{1/2}\}} \right] \xrightarrow{d} 0 \quad (P_0).$$

Die Beziehungen (8) bis (10) besagen genau, dass die Voraussetzungen des Brownschen Satzes erfüllt sind; deshalb folgt (3). ■

Die Voraussetzung (1) des Satzes 1 ist in wichtigen Spezialfällen erfüllt. Es gilt:

2 Korollar

Sei $p_0 \in p(r)$ und $P_{1,0}$ als stationäre Verteilung von p_0 eindeutig bestimmt. Dann gilt für jede Kurve $s \rightarrow p_s \in p(r)$, die $L^2(P_{2,0})$ -differenzierbar mit Tangentialvektor g ist, dass für jede Folge $(u_n : n \in \mathbb{N}) \subset \mathbb{R} \setminus \{0\}$ mit $u_n \rightarrow u'$

$$\frac{dp_{n-1, u_n n^{-1/2}}}{dp_{n-1, 0}} =$$

$$\exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} g(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}(\xi^2) + o_{P_{n,0}}(n^0) \right]$$

und

$$\left[n^{-1/2} \sum_{i=1}^{n-1} g(x_i, x_{i+1}) \right] * P_{n,0} \longrightarrow N(0, P_{2,0}(\xi^2)).$$

Beweis

Die Eindeutigkeit der stationären Verteilung $P_{1,0}$ von p_0 hat die Ergodizität des Markoff-Prozesses $\left[(E^{\mathbb{N}}, \mathcal{B}(E^{\mathbb{N}}), P_0, (X_i : i \in \mathbb{N})) \right]$ zur Folge. Einen Beweis dieser Aussage findet man beispielsweise in Billingsley (1), p. 52. Birkhoffs Ergodensatz besagt dann gerade die Gültigkeit der Beziehung (1) aus Satz 1. Damit ergibt sich die Behauptung aus Satz 1. ■

Mit Hilfe des Satzes 1 können wir nun folgendes hinreichende Kriterium für die lokale asymptotische Normalität der Folge der statistischen Experimente

$$F_n := \left[E^n, \mathcal{B}(E^n), \left\{ P_{n,h} : h \in H \right\} \right]$$

angeben.

3 Satz

Es sei h' aus dem topologischen Inneren von H . $P_{1,h'}$ bezeichne eine stationäre Verteilung von $p_{h'} \in p(H)$. Die Familie $p(H)$ sei in $h' L^2(P_{2,h'})$ -differenzierbar mit injektiver Ableitung

$$Dp_{h'} : H \longrightarrow L^2(P_{2,h'})$$

(vgl. 5.13). Für jedes $h \in H$ gelte

$$(1) \quad \frac{1}{n} \sum_{i=1}^n p_{h'} \left[x_i \left(Dp_{h'}(h)(x_{i,n}) \right)^2 \right] \longrightarrow P_{2,h'} \left[\left(Dp_{h'}(h) \right)^2 \right]$$

$P_{n,h'}$ -stochastisch.

Dann ist für jede Familie von Startverteilungen $\left\{ P_{1,h} : h \in H \right\}$ mit

$$(2) \quad \frac{dP_{1, h}}{dP_{1, h'}} \longrightarrow 1 \quad (P_{1, h'}) \text{ für } |h - h'| \longrightarrow 0$$

und jede Folge von Elementen h_n aus H mit

$$(3) \quad h_n \longrightarrow h \in H$$

folgendes wahr.

$$(4) \quad \frac{dP_{n, h' + h_n n^{-1/2}}}{dP_{n, h'}} = \exp \left[L_n(h) - \frac{1}{2} \|h\|_{h'}^2 + o_{P_{n, h'}}(n^0) \right].$$

Dabei bezeichnet $\|\cdot\|_{h'}$ die Norm, die von dem Skalarprodukt

$$(5) \quad \langle \cdot, \cdot \rangle_{h'} := P_{2, h'} \left[Dp_{h'}(\cdot) Dp_{h'}(\cdot) \right]$$

herrührt. L_n ist für jedes $n \in \mathbb{N}$ der lineare Prozess

$$(6) \quad L_n(h)(\underline{x}) := n^{-1/2} \sum_{i=1}^{n-1} Dp_{h'}(h)(x_i, x_{i+1}).$$

Es gilt

$$(7) \quad L_n(h) * P_{n, h'} \longrightarrow N(0, \|h\|_{h'}^2) \quad \text{schwach.}$$

Beweis

Zunächst stellen wir fest, dass

$$\begin{aligned} & \frac{dP_{n, h' + h_n n^{-1/2}}}{dP_{n, h'}} = \\ & \frac{dP_{1, h' + h_n n^{-1/2}}}{dP_{1, h'}} \frac{dp_{n-1, h' + h_n n^{-1/2}}}{dp_{n-1, h'}}. \end{aligned}$$

Wegen (2) ist

$$\frac{dP_{1, h' + h_n n^{-1/2}}}{dP_{1, h'}} = \exp \left[o_{P_{n, h'}}(n^0) \right].$$

Aus der Differenzierbarkeit der Familie $p(H)$ folgt laut 6.16 die $L^2(P_{2, h'})$ -Differenzierbarkeit jeder der Kurven

$$s \longrightarrow p_{h' + s h_g},$$

wenn $h_s \rightarrow h$ für ein $h \in H$. Der entsprechende Tangentialvektor ist dann $Dp_{h'}(h)$. Wähle eine Kurve $s \rightarrow h_s$ mit $h_{s^{-1/2}} = h_n$. Die

Voraussetzung (1) ergibt hiermit zusammen die Gültigkeit der Forderungen aus 1. Deshalb erhalten wir

$$\frac{dp_{n^{-1}, h' + h_n^{-1/2}}}{dp_{n^{-1}, h'}}(\underline{x}) =$$

$$\exp \left[n^{-1/2} \sum_{i=1}^{n-1} Dp_{h'}(h) \langle x_i, x_{i+1} \rangle - \right.$$

$$\left. \frac{1}{2} P_{z, h'} \left[\left(Dp_{h'}(h) \right)^2 \right] + o_{P_{n, h'}}(n^0) \right].$$

Offenbar ist L_n wie unter (6) linear und aus $L^2(P_{n, h'})$. Da die Ableitung $Dp_{h'}$ injektiv ist laut unseren Voraussetzungen, ist durch (5) ein Skalarprodukt definiert, für welches

$$P_{z, h'} \left[\left(Dp_{h'}(h) \right)^2 \right] = \|h\|_{h'}^2$$

gilt. Die Aussage (7) ergibt sich aus 1.(3) und (5). ■

4 Korollar

Unter den Voraussetzungen des Satzes 3 ist die Folge der statistischen Experimente

$$F_n := \left[E^n, \mathcal{B}(E^n), \left\{ P_{n, h} : h \in H \right\} \right]$$

in h' lokal asymptotisch normal. Genauer gilt: Die Folge der in h' lokalisierten Experimente

$$F_{n, h'} := \left[E^n, \mathcal{B}(E^n), \left\{ P_{n, h' + n^{-1/2} h} : h \in H_n \right\} \right]$$

mit

$$H_n := \left\{ h \in H : h' + n^{-1/2} h \in H \right\}$$

konvergiert schwach gegen den Gauss-Shift auf $\left[H, \langle \cdot, \cdot \rangle_{h'} \right]$.

Beweis

Strasser(1), 80.2 ■

5 Korollar

Die Voraussetzungen von 3 seien erfüllt bis auf (1). Statt dessen sei die Verteilung $P_{1,h'}$ durch die Eigenschaft stationär für $p_{h'}$ zu sein eindeutig bestimmt. Dann gilt die Aussage des Satzes 3 unvermindert und ebenso die des Korollars 4.

Beweis

Ganz entsprechend zu Korollar 2 sieht man diese Aussage. ■

6 Bemerkung

Wir weisen noch darauf hin, dass es für die Gültigkeit der Korollare 2 und 5 letztlich nicht auf die Eindeutigkeit der stationären Verteilung, sondern die Ergodizität der Markoff-Prozesse ankommt. ■

Die Aussage des Satzes 3 soll nun auf den Fall zeitlich homogener Markoff-Prozesse ausgedehnt werden. Wir fragen also nach der lokalen asymptotischen Normalität, wenn die Startverteilung nicht stationär ist.

7 Satz

Sei h' ein innerer Punkt von H und die stationäre Verteilung $P_{1,h'}$ von $p_{h'}$ sei eindeutig bestimmt. Wir nehmen an, die Familie $p(H)$ sei in h' $L^2(P_{2,h'})$ -differenzierbar und die Ableitung

$$Dp_{h'} : H \longrightarrow L^2(P_{2,h'})$$

sei injektiv. Ferner gelte:

$$(1) \quad h \longmapsto \frac{dp_h}{dp_{h'}} \text{ ist stetig an der Stelle } h'$$

und

$$(2) \quad p_{h'}(x_n) \ll P_{1,h'} \text{ für alle } x \in E.$$

Sind dann $\bar{P}_{1,h}$ beliebige Startverteilungen mit

$$(3) \quad \frac{d\bar{P}_{1,h}}{dP_{1,h'}} \longrightarrow 1 \quad (P_{1,h'})$$

und h_n Elemente aus H mit $h_n \longrightarrow h \in H$, so gilt

$$(4) \quad \frac{d\bar{P}_{n, h'} + h_n^{-1/2}}{d\bar{P}_{n, h'}} = \exp \left[L_n(h) - \frac{1}{2} \|h\|_{h'}^2 + o_{P_{n, h'}}^{n^0} \right].$$

Dabei ist $\|\cdot\|_{h'}^2 = \langle \cdot, \cdot \rangle_{h'}$ mit dem Skalarprodukt

$$(5) \quad \langle \cdot, \cdot \rangle_{h'} := P_{2, h'} \left[Dp_{h'}(\cdot), Dp_{h'}(\cdot) \right].$$

L_n ist für jedes $n \in \mathbb{N}$ der lineare Prozess

$$(6) \quad L_n(h)(\underline{x}) := n^{-1/2} \sum_{i=1}^{n-1} Dp_{h'}(h)(x_i, x_{i+1}).$$

Es gilt

$$(7) \quad L_n(h) \star \bar{P}_{n, h'} \longrightarrow N(0, \|h\|_{h'}^2) \quad \text{schwach.}$$

Beweis

Die Voraussetzungen des Korollars 2 sind erfüllt, weswegen

$$\begin{aligned} & \frac{dp_{n-1, h'} + h_n^{-1/2}}{dp_{n-1, h'}}(\underline{x}) = \\ & \exp \left[n^{-1/2} \sum_{i=1}^{n-1} Dp_{h'}(h)(x_i, x_{i+1}) - \right. \\ & \quad \left. \frac{1}{2} P_{2, h'} \left[\left(Dp_{h'}(h) \right)^2 \right] + o_{P_{n, h'}}^{n^0} \right] = \\ & \exp \left[L_n(h) - \frac{1}{2} \|h\|_{h'}^2 + o_{P_{n, h'}}^{n^0} \right]. \end{aligned}$$

Der Rest $o_{P_{n, h'}}^{n^0}$ erfüllt also

$$(8) \quad o_{P_{n, h'}}^{n^0} = \log \frac{dp_{n-1, h'} + h_n^{-1/2}}{dp_{n-1, h'}} - L_n(h) + \frac{1}{2} \|h\|_{h'}^2 = \sum_{i=1}^{n-1} \left\{ \log \frac{dp_{h'} + h_n^{-1/2}}{dp_{h'}}(x_i, x_{i+1}) - n^{-1/2} Dp_{h'}(h)(x_i, x_{i+1}) \right\}$$

$$+ \frac{1}{2n-2} \|h\|_{h'}^2 \Big\}.$$

Die Summanden der rechten Seite von (8) streben gemäss (1) gegen 0. Darum sind die Voraussetzungen des Lemmas 8.5 erfüllt, und wir können schliessen, dass die Reste $o_p(n^0)$ auch in $\bar{P}_{n,h'}$ -Wahrscheinlichkeit verschwinden.

Die Konvergenz unter (7) ist noch nachzuweisen. Sie ergibt sich, wenn man Lemma 8.5 auf die Voraussetzungen des Satzes von Brown anwendet und wie in Satz 3 verfährt. ■

8 Korollar

Unter den Voraussetzungen des Satzes 7 ist die Folge der statistischen Experimente

$$F_n := \left[E^n, \mathcal{B}(E^n), \left\{ \bar{P}_{n,h} : h \in H \right\} \right]$$

in h' lokal asymptotisch normal. Genauer gilt: Die Folge der in h' lokalisierten Experimente

$$F_{n,h'} := \left[E^n, \mathcal{B}(E^n), \left\{ \bar{P}_{n,h'+nh}^{-1/2} : h \in H_n \right\} \right]$$

mit

$$H_n := \left\{ h \in H : h' + n^{-1/2}h \in H \right\}$$

konvergiert schwach gegen den Gauss-Shift auf $\left[H, \langle \cdot, \cdot \rangle_{h'} \right]$. Das lokale Grenzexperiment hängt also nicht von den Startverteilungen ab, sofern diese nur stetig im Sinne von 7 (3) sind.

Beweis

Strasser (1), 80.2 ■

9 Bemerkung

Betrachte ein σ -finites Mass R auf einem messbaren Raum $(\Omega, \mathcal{A})$. $P(R)$ bezeichne die Familie der bzgl. R absolut stetigen Verteilungen auf $(\Omega, \mathcal{A})$. Für $P \in P(R)$ sei P^n das n -fache Produkt der Verteilung P mit sich selbst. Dann besteht ein enger Zusammenhang zwischen der Gültigkeit stochastischer Entwicklungen und der Differenzierbarkeit von Kurven in $P(R)$. Ist $g \in L^2(P_0)$ mit $P_0(g) = 0$, so treffen folgende zwei Aussagen zu.

(1) Wenn die Kurve $s \rightarrow P_s \in P(R)$ differenzierbar im Sinne der

Definition 75.1 aus Strasser (1) ist, dann gilt für jede Folge u_n reeller Zahlen mit $u_n \rightarrow u_0$

$$\frac{d \left(P_{u_n^{-1/2}} \right)^n}{dP_0^n} (\omega_1, \dots, \omega_n) = \exp \left[u_n^{-1/2} \sum_{i=1}^n \xi(\omega_i) - \frac{u_n^2}{2} P_0(\xi^2) + o_{P_0^n}(n^0) \right].$$

(2) Ist für jede Folge reeller Zahlen u_n mit $|u_n| \rightarrow 1$

$$\frac{d \left(P_{u_n^{-1/2}} \right)^n}{dP_0^n} (\omega_1, \dots, \omega_n) = \exp \left[u_n^{-1/2} \sum_{i=1}^n \xi(\omega_i) - \frac{u_n^2}{2} P_0(\xi^2) + o_{P_0^n}(n^0) \right],$$

so ist die Kurve $s \rightarrow P_s$ differenzierbar im Sinne der Definition 75.1, Strasser (1) mit Tangentialvektor ξ .

Aussagen dieser Art findet man beispielsweise in Strasser (1), 76. oder auch in Pflanzagl (2), 1.

In den vorangegangenen Sätzen haben wir dargelegt, dass die $L^2(P_{2,0})$ -Differenzierbarkeit einer Kurve von Markoff-Kernen aus $p(r)$ die Gültigkeit einer stochastischen Entwicklung wie in 1 oder in 2 impliziert. Umgekehrt fragen wir nun nach der Differenzierbarkeit einer solchen Kurve unter der Annahme, eine solche stochastische Entwicklung sei zulässig.

10 Diskussion

Betrachte einen Markoff-Kern p_0 aus $p(r)$, dessen stationäre Verteilung eindeutig bestimmt ist. p_s sollen weitere Markoff-Kerne aus $p(r)$ sein, und es gelte für jede Folge reeller Zahlen $(u_n : n \in \mathbb{N})$, die gegen eine Zahl $u_0 \in \mathbb{R}$ konvergiert

$$(1) \quad \frac{dp_{n-1, u_n^{-1/2}}}{dp_{n-1, 0}} =$$

$$\exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} g(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0} \left(g^2 \right) + \sum_{i=1}^{n-1} r_{u_n n}^{-1/2} \right]$$

mit

$$(2) \quad \sum_{i=1}^{n-1} r_{u_n n}^{-1/2} = o_{P_{n,0}}(1)$$

und

$$(3) \quad g \in L^2(P_{2,0}) \text{ und } P_{2,0}(g) = 0.$$

Aus (3) folgt die Existenz einer $L^2(P_{2,0})$ -differenzierbaren Kurve $s \rightarrow q_s$ in $p(r)$ mit Tangentialvektor g (vgl. 6.4 und 6.9). Dabei ist $q_0 = p_0$. Aufgrund der Eindeutigkeit der stationären Verteilung $P_{1,0}$ von q_0 ergibt sich mit Hilfe des Korollars 2, dass

$$(4) \quad \frac{dq_{n-1, u_n n}^{-1/2}}{dq_{n-1, 0}} =$$

$$\exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} g(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0} \left(g^2 \right) + \sum_{i=1}^{n-1} t_{u_n n}^{-1/2} \right]$$

mit

$$(5) \quad \sum_{i=1}^{n-1} t_{u_n n}^{-1/2} = o_{P_{n,0}}(1)$$

und

$$(6) \quad \left[n^{-1/2} \sum_{i=1}^{n-1} g(x_i, x_{i+1}) \right] * P_{n,0} \rightarrow N(0, P_{2,0}(g^2)).$$

Der Hellinger-Abstand der Verteilungen $P_{1,0} \otimes p_{n-1, u_n n}^{-1/2}$ und

$P_{1,0} \otimes p_{n-1, u_n n}^{-1/2}$ konvergiert gegen 0. Bezeichnet nämlich Y eine

$N(0, P_{2,0}(g^2))$ -verteilte Zufallsvariable, so ist wegen (2)

$$(7) \quad \frac{dP_{1,0}^{\otimes p}}{dP_{n,0}} \Big|_{n^{-1}, u_n^{-1/2}} =$$

$$\exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}(\xi^2) + \sum_{i=1}^{n-1} r_{u_n^{-1/2}} \right] \longrightarrow$$

$$\exp \left[u_0 Y - \frac{u_0^2}{2} P_{2,0}(\xi^2) \right] \quad \text{schwach unter } P_{n,0}.$$

Mit Hilfe des Lemmas 6.12 aus Strasser (1) erhalten wir hieraus die gleichgradige $P_{n,0}$ -Integrierbarkeit der Folge

$$(8) \quad \left[\frac{dP_{1,0}^{\otimes p}}{dP_{n,0}} \Big|_{n^{-1}, u_n^{-1/2}} : n \in \mathbb{N} \right].$$

Ebenso weist man die gleichgradige $P_{n,0}$ -Integrierbarkeit der Folge

$$(9) \quad \left[\frac{dP_{1,0}^{\otimes q}}{dP_{n,0}} \Big|_{n^{-1}, u_n^{-1/2}} : n \in \mathbb{N} \right]$$

nach. Die Differenzen

$$\frac{dP_{1,0}^{\otimes p}}{dP_{n,0}} \Big|_{n^{-1}, u_n^{-1/2}} - \frac{dP_{1,0}^{\otimes q}}{dP_{n,0}} \Big|_{n^{-1}, u_n^{-1/2}}$$

streben offenbar $P_{n,0}$ -stochastisch gegen 0; wegen (8) und (9) also auch im $P_{n,0}$ -Mittel. Anhand der Gleichung (7) sehen wir, dass die beiden binären Experimentenfolgen

$$\left[E^n, \mathcal{B}(E^n), \left[P_{n,0}, P_{1,0}^{\otimes p} \Big|_{n^{-1}, u_n^{-1/2}} \right] \right]$$

und

$$\left[E^n, \mathcal{B}(E^n), \left[P_{n,0}, P_{1,0}^{\otimes q} \Big|_{n^{-1}, u_n^{-1/2}} \right] \right]$$

schwach konvergieren. Deshalb sind die Folgen (8) und (9) benachbart. Hieraus folgt nun

$$(10) \quad d_1 \left[P_{1,0} \otimes p_{n-1, u_n}^{-1/2}, P_{1,0} \otimes q_{n-1, u_n}^{-1/2} \right] \longrightarrow 0.$$

Die Äquivalenz von Hellinger-Abstand und Totalvariation (vgl. 11.2.(7)) sowie die Beziehung 11.2.(3) für den Hellinger-Abstand und die Affinität α zweier Verteilungen ergeben dann

$$(11) \quad \alpha \left[P_{1,0} \otimes p_{n-1, u_n}^{-1/2}, P_{1,0} \otimes q_{n-1, u_n}^{-1/2} \right] \longrightarrow 1.$$

Die Umkehrung dieser Aussage ist ebenfalls richtig, präziser haben wir:

Ist $s \longrightarrow q_s$ eine $L^2(P_{2,0})$ -differenzierbare Kurve in $p(r)$ mit Tangentialvektor ξ und gilt für die Kurve $s \longrightarrow p_s$ mit

$q_0 = p_0$ die Beziehung (11), so sind (1) und (2) erfüllt und natürlich ebenso (3) und (6). Die letzten beiden Aussagen folgen einfach aus der Tatsache, dass ξ Tangentialvektor einer differenzierbaren Kurve ist. Um (1) und (2) nachzuweisen, wähle $r_{u_n}^{-1/2}$ derart, dass

$$\frac{dp_{u_n}^{-1/2}}{dp_0} = \exp \left[n^{-1/2} u_n \xi - \frac{u_n^2}{2n-2} P_{2,0}(\xi^2) + r_{u_n}^{-1/2} \right].$$

Dann gilt wegen (11)

$$P_{n,0} \left[\left| \frac{dp_{n-1, u_n}^{-1/2}}{dp_{n-1,0}} - \frac{dq_{n-1, u_n}^{-1/2}}{dp_{n-1,0}} \right| \right] =$$

$$P_{1,0} \otimes r_{n-1} \left[\left| \frac{dp_{n-1, u_n}^{-1/2}}{dr_{n-1}} - \frac{dq_{n-1, u_n}^{-1/2}}{dr_{n-1}} \right| \cdot 1_{\left\{ \frac{dp_{n-1,0}}{dr_{n-1}} \neq 0 \right\}} \right]$$

$\longrightarrow 0$. Da nun

$$\frac{dq_{n-1, u_n}^{-1/2}}{dq_{n-1,0}}(\underline{x}) =$$

$$\exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}(\xi^2) + \sum_{i=1}^{n-1} t_{u_n}^{-1/2}(x_i, x_{i+1}) \right],$$

folgt

$$\begin{aligned} & \exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}(\xi^2) + \sum_{i=1}^{n-1} r_{u_n}^{-1/2}(x_i, x_{i+1}) \right] - \\ & \exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}(\xi^2) + \sum_{i=1}^{n-1} t_{u_n}^{-1/2}(x_i, x_{i+1}) \right] \longrightarrow 0 \text{ im } P_{n,0}\text{-Mittel.} \end{aligned}$$

Daher ist

$$\frac{dq_{n-1, u_n}^{-1/2}}{dp_{n-1, 0}} \left[1 - \exp \left\{ \sum_{i=1}^{n-1} r_{u_n}^{-1/2} - \sum_{i=1}^{n-1} t_{u_n}^{-1/2} \right\} \right] \longrightarrow 0 \quad (P_{n,0}),$$

und

$$\exp \left\{ \sum_{i=1}^{n-1} r_{u_n}^{-1/2} - \sum_{i=1}^{n-1} t_{u_n}^{-1/2} \right\} \longrightarrow 1 \quad (P_{n,0}).$$

Weil

$$\sum_{i=1}^{n-1} t_{u_n}^{-1/2} \longrightarrow 0 \quad (P_{n,0}),$$

folgt hieraus

$$\sum_{i=1}^{n-1} r_{u_n}^{-1/2} \longrightarrow 0 \quad (P_{n,0}).$$

Dies ergibt (2) und (1). Wir haben also folgendes Resultat.

11 Satz

Ist $\xi \in L^2(P_{2,0})$ und $p_0 \langle x, \xi(x, \cdot) \rangle = 0$, so gilt für eine beliebige Kurve $s \longrightarrow p_s \in p(r)$ und jede Folge $\langle u_n : n \in \mathbb{N} \rangle \subset \mathbb{R}$ mit $u_n \longrightarrow u_0$

$$\frac{dp_{n-1, u_n^{-1/2}}}{dp_{n-1, 0}} = \exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}(\xi^2) + o_{P_{n,0}}(1) \right]$$

dann und nur dann, wenn eine $L^2(P_{2,0})$ -differenzierbare Kurve $s \longrightarrow q_s \in p(r)$ existiert mit Tangentialvektor ξ und $q_0 = p_0$, so dass

$$\alpha \left(P_{1,0} \otimes p_{n-1, u_n^{-1/2}}, P_{1,0} \otimes q_{n-1, u_n^{-1/2}} \right) \longrightarrow 1.$$

Es gilt dann

$$\left[n^{-1/2} \sum_{i=1}^{n-1} \xi(x_i, x_{i+1}) \right] * P_{n,0} \longrightarrow N(0, P_{2,0}(\xi^2)).$$

■

12 Diskussion

Beachte nun die Aussage 6.11, die besagt, dass in der Situation des Satzes 11 die Kurve $s \longrightarrow p_s$ dann und nur dann $L^2(P_{2,0})$ -differenzierbar ist, wenn

$$(1) \quad \alpha \left(P_{1,0} \otimes p_s, P_{1,0} \otimes q_s \right) \left[|s|^{-2} \right] \longrightarrow 1 \text{ für } s \longrightarrow 0.$$

Die Gleichung (1) ist ihrerseits gleichwertig mit

$$(2) \quad \alpha \left(P_{1,0} \otimes p_{u_n^{-1/2}}, P_{1,0} \otimes q_{u_n^{-1/2}} \right)^n \longrightarrow 1$$

für beliebige Folgen reeller Zahlen u_n mit $|u_n| \longrightarrow 1$. Dies folgt beispielsweise mit Hilfe des Lemmas 11.6.

An dieser Stelle treten die Schwierigkeiten deutlich zu Tage, die bei dem Versuch auftreten, eine zu 9 analoge Aussage zu treffen. Nehmen wir zunächst an, die Kerne q_s und p_s seien Verteilungen auf

$\langle E, \mathfrak{B}(E) \rangle$, so gilt

$$p_{n-1, u_n}^{-1/2} = \left(p_{u_n}^{-1/2} \right)^{n-1}$$

$$q_{n-1, u_n}^{-1/2} = \left(q_{u_n}^{-1/2} \right)^{n-1}$$

und deshalb (vgl. 7.2.(9))

$$(3) \quad \alpha \left(P_{1,0} \otimes p_{n-1, u_n}^{-1/2}, P_{1,0} \otimes q_{n-1, u_n}^{-1/2} \right) =$$

$$\alpha \left(\left(p_{u_n}^{-1/2} \right)^{n-1}, \left(q_{u_n}^{-1/2} \right)^{n-1} \right) =$$

$$\alpha \left(p_{u_n}^{-1/2}, q_{u_n}^{-1/2} \right)^{n-1}.$$

In diesem Fall sind also eine stochastische Entwicklung wie unter 10 (1) - (3) und die $L^2(P_{2,0})$ -Differenzierbarkeit der Kurve äquivalent. Im allgemeineren Fall echter Übergangskerne ist die Gleichung (3) nicht zulässig, denn sie beruht wesentlich auf der Produktmasseigenschaft der beteiligten Masse, also auf der Unabhängigkeit der Versuchswiederholungen der zugrundeliegenden Experimentenfolge.

Die gleichen Schwierigkeiten treten übrigens auf, wenn man versucht im Markoff-Fall auf der Grundlage des Pfanzaglschen DCC-Differenzierbarkeitsbegriffs zu arbeiten. Grob gesprochen sind ja hier für eine Kurve $s \rightarrow p_s$ Bedingungen anzugeben für die Gültigkeit von

$$\frac{dp_s}{dp_0} = 1 + sg + r_s$$

mit

$$(4) \quad \sum_{i=1}^{n-1} r_{u_n}^{-1/2} (x_i, x_{i+1}) \rightarrow 0 \quad \text{schwach}$$

für Folgen $(u_n : n \in \mathbb{N})$ mit $u_n \rightarrow u_0$.

Im i.i.d. Fall, wenn also die Kerne p_s Verteilungen sind, gibt das DCC-Kriterium (vgl. Loeve, p.329) äquivalente Bedingungen für die schwache Konvergenz der Restesummen in (4). Diese Bedingungen entsprechen im i.i.d. Fall gerade der Definition der

DCC-Differenzierbarkeit. Die Beweise des DCC-Kriteriums, wie sie ursprünglich von Kolmogoroff, 1928, und später mit Hilfe der Fourier-Analyse präsentiert wurden, bauen aber wesentlich auf der Unabhängigkeitsvoraussetzung auf. Eine hinreichende und notwendige DCC-Bedingung scheint für den Markoff-Fall nicht bekannt zu sein.

Bei dem Versuch der Konstruktion eines Beispiels einer Kurve $s \rightarrow p_s \in p(r)$, die 10 (1) - (3) erfüllt, aber nicht differenzierbar ist, tritt als Problem ebenfalls die mangelnde Kenntnis aquivalenter Bedingungen für die degenerierte Konvergenz von Folgen der Form

$$\left(\sum_{l=1}^{n-1} r_n(x_l, x_{l+1}) : n \in \mathbb{N} \right)$$

unter den Randverteilungen P_n eines Markoff-Prozesses auf. Sei nämlich p_0 ein beliebiger Markoff-Kern mit stationärer Verteilung $P_{1,0}$ und $\xi \in L^2(P_{2,0})$ mit $p_0(x, \xi(x)) = 0$. Ferner sei r_n definiert durch

$$\frac{dp_{u_n n^{-1/2}}}{dp_0} = \exp \left[n^{-1/2} u_n \xi - \frac{u_n^2}{2n-2} P_{2,0}(\xi^2) + r_{u_n n^{-1/2}} \right].$$

Damit 10 (1) - (3) erfüllt sind, muss

$$(5) \quad \sum_{l=1}^{n-1} r_{u_n n^{-1/2}}(x_l, x_{l+1}) \rightarrow 0 \quad (P_{n,0})$$

gelten. Das Problem ist nun, einen Rest zu finden, welcher (5) genügt, ohne dass die Kurve $s \rightarrow p_s$ bereits differenzierbar ist. Schon in einfachen Fällen scheint uns die Situation unüberschaubar. Andererseits kann man aus der Gültigkeit von (1) - (3) und gewissen zusätzlichen Annahmen die Differenzierbarkeit der Kurve $s \rightarrow p_s$ schliessen. Wir geben einige Beispiele.

14 Satz

Für eine Kurve $s \rightarrow p_s$ sollen 10 (1) - (3) erfüllt sein. Gilt für die Reste $r_{u_n n^{-1/2}}$ zusätzlich

$$(1) \quad n \cdot P_{2,0} \left[\left| r_{u_n n^{-1/2}} \right| \right] \longrightarrow 0,$$

so ist die Kurve $L^2(P_{2,0})$ -differenzierbar mit Tangentialvektor g .

Beweis

Diese Aussage folgt aus dem folgenden Satz 15; dabei ist nur zu beachten, dass (1)

$$\sum_{i=1}^n r_{u_n n^{-1/2}}(x_i, y_i) \longrightarrow 0 \quad \text{schwach}$$

unter $(P_{2,0})^n$ impliziert. ■

15 Satz

Sei $g \in L^2(P_{2,0})$ und $p_0(x, g(x, \cdot)) = 0$. Genau dann gilt für jede Folge reeller Zahlen u_n mit einem Grenzwert u_0 , dass

$$(1) \quad \frac{dP_{1,0}^{\otimes p_{n-1, u_n n^{-1/2}}}}{dP_{n,0}} = \exp \left[n^{-1/2} u_n \sum_{i=1}^{n-1} g(x_i, x_{i+1}) - \frac{u_n^2}{2} P_{2,0}(g^2) + \sum_{i=1}^{n-1} r_{u_n n^{-1/2}}(x_i, x_{i+1}) \right]$$

und

$$(2) \quad \sum_{i=1}^{n-1} r_{u_n n^{-1/2}}(x_i, y_i) \longrightarrow 0 \quad (P_{2,0}^n),$$

wenn die Kurve $s \longrightarrow p_s$ $L^2(P_{2,0})$ -differenzierbar mit Tangentialvektor g ist.

Beweis

Es gelte (1). Dann ist

$$\frac{dp_{u_n n^{-1/2}}}{dp_0}(x, y) =$$

$$\exp \left[n^{-1/2} u_n \xi(x,y) - \frac{u_n^2}{2n-2} P_{2,0} \left(\xi^2 \right) + r_{u_n n^{-1/2}}(x,y) \right]$$

Also folgt

$$(3) \quad \frac{d_{i=1}^n P_{1,0} \otimes P_{u_n n^{-1/2}}}{dP_{2,0}^n} \left[(x_1, y_1), \dots, (x_n, y_n) \right] =$$

$$\exp \left[n^{-1/2} u_n \sum_{i=1}^n \xi(x_i, y_i) - \frac{u_n^2}{2} P_{2,0} \left(\xi^2 \right) + u_n^2 \left(\frac{1}{2} - \frac{n}{2n-2} \right) P_{2,0} \left(\xi^2 \right) + \sum_{i=1}^n r_{u_n n^{-1/2}}(x_i, y_i) \right].$$

Definiere

$$\hat{r}_{u_n n^{-1/2}}(x_i, y_i) :=$$

$$n^{-1} u_n^2 \left(\frac{1}{2} - \frac{n}{2n-2} \right) P_{2,0} \left(\xi^2 \right) + r_{u_n n^{-1/2}}(x_i, y_i).$$

Wegen (2) erhalten wir hieraus

$$(4) \quad \sum_{i=1}^n \hat{r}_{u_n n^{-1/2}}(x_i, y_i) \longrightarrow 0 \quad (P_{2,0}^n).$$

Gemäss Strasser (1), 75.9 folgt nun

$$(5) \quad n u_n^{-2} P_{1,0} \otimes r \left[\left\{ 2 \left[\frac{d P_{1,0} \otimes P_{u_n n^{-1/2}}}{dP_{1,0} \otimes r} \right]^{1/2} - 2 \left[\frac{dP_{2,0}}{dP_{1,0} \otimes r} \right]^{1/2} - u_n n^{-1/2} \xi \left[\frac{dP_{2,0}}{dP_{1,0} \otimes r} \right]^{1/2} \right\}^2 \right] \longrightarrow 0.$$

Mit Hilfe von 11.6 schliessen wir hieraus die $L^2(P_{1,0} \otimes r)$ -Differenzierbarkeit mit Tangentialvektor ξ der Kurve $s \longrightarrow p_s$ und weiter gemäss 6.9 die erste Teilbehauptung.

Umgekehrt nehmen wir jetzt an, die Kurve sei $L^2(P_{2,0})$ -differenzierbar. Dann ist laut Korollar 2 die Gleichung (1) gültig. Die Kurve von Verteilungen, $s \longrightarrow P_{1,0} \otimes p_s$, ist

differenzierbar im Sinne der Definition 75.1, Strasser (1). Nach 75.8, ebenda, ist dann

$$(5) \quad \frac{d \otimes_{i=1}^n P_{1,0} \otimes_{u_n^{-1/2}}}{dP_{2,0}^n} \left(\langle x_1, y_1 \rangle, \dots, \langle x_n, y_n \rangle \right) =$$

$$\exp \left(n^{-1/2} u_n \sum_{i=1}^n \xi \langle x_i, y_i \rangle - \frac{u_n^2}{2} P_{2,0} \left(\xi^2 \right) + \right.$$

$$\left. \sum_{i=1}^n \hat{r}_{u_n^{-1/2}} \langle x_i, y_i \rangle \right)$$

mit

$$(6) \quad \sum_{i=1}^n \hat{r}_{u_n^{-1/2}} \langle x_i, y_i \rangle \longrightarrow 0 \quad (P_{2,0}^n) .$$

Aus (1), (5) und (6) folgt

$$\log \frac{dp_{u_n^{-1/2}}}{dp_0} =$$

$$n^{-1/2} u_n \xi - \frac{u_n^2}{2n-2} P_{2,0} \left(\xi^2 \right) + r_{u_n^{-1/2}} =$$

$$n^{-1/2} u_n \xi - \frac{u_n^2}{2n} P_{2,0} \left(\xi^2 \right) + \hat{r}_{u_n^{-1/2}} .$$

Hieraus folgern wir

$$\sum_{i=1}^n \hat{r}_{u_n^{-1/2}} \langle x_i, y_i \rangle \longrightarrow 0 \quad (P_{2,0}^n) ,$$

also die Beziehung (2). ■

3. Lokal schwache Konvergenz von Experimenten - differenzierbare Mannigfaltigkeiten als Parameterraume

3.1 Lokal schwache Konvergenz

Sei M eine differenzierbare Hilbert-Mannigfaltigkeit in einem Banach-Raum B . Den zugrundeliegenden Hilbert-Raum bezeichnen wir mit H . Er sei separabel. Der hier verwendete Begriff einer differenzierbaren Mannigfaltigkeit weicht von dem üblicherweise in der Literatur zu findenden ab, vgl. hierzu den Abschnitt 10. Den Tangentialraum an M in $m' \in M$ notieren wir mit $T_{g_{m'}}M$. Er ist stets ein Hilbert-Raum.

$\langle \cdot, \cdot \rangle_{m'}$ steht für ein Skalarprodukt auf $T_{g_{m'}}M$, das von m' abhängen darf. Die dadurch gegebene Norm bezeichnen wir mit $\|\cdot\|_{m'}$. $\left\{ \Omega_n, \mathcal{A}_n \right\}$ sei eine beliebige Folge messbarer Räume. Wir betrachten eine Folge statistischer Experimente

$$F_n := \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right\}.$$

Das Konzept der lokalen asymptotischen Normalität soll auf diesen Fall übertragen werden.

1 Definition

Die Folge der Experimente F_n konvergiert in m' lokal schwach gegen einen Gauss-Shift für $H = \left\{ T_{g_{m'}}M, \langle \cdot, \cdot \rangle_{m'} \right\}$, falls für jede Folge von Elementen m_n aus M mit

$$n^{1/2}(m_n - m') \longrightarrow t \quad (t \in T_{g_{m'}}M)$$

gilt

$$\frac{dP_{n,m_n}}{dP_{n,m'}} = \exp \left[L_n(t) - \frac{1}{2} \|t\|_{m'}^2 + o_{P_{n,m'}}(n^0) \right].$$

Dabei ist $(L_n(t) : t \in T_{g_{m'}}M)$ für jedes $n \in \mathbb{N}$ ein linearer Prozess, der

$$L_n(t) * P_{n,m'} \longrightarrow N(0, \|t\|_{m'}^2) \quad \text{schwach für } n \longrightarrow \infty$$

für alle $t \in T_{g_{m'}}M$ erfüllt. $L_n(t)$ heisst Folge der lokal zentralen Prozesse für F_n . Vereinfachend sagen wir auch, die Folge der Experimente F_n ist in m' lokal asymptotisch normal. ■

2 Bemerkung und Definition

In der Situation der Definition 1 sei speziell $T_{\mathcal{E}_{m'}} M$ ein euklidischer Raum, also von endlicher Dimension. Dann existiert eine messbare Abbildung

$$X_n : \left(\Omega_n, \mathcal{A}_n \right) \longrightarrow \left(T_{\mathcal{E}_{m'}} M, \mathcal{B}(T_{\mathcal{E}_{m'}} M) \right)$$

für jedes $n \in \mathbb{N}$ mit den Eigenschaften

$$(1) \quad L_n(t)(\omega_n) = \langle t, X_n(\omega_n) \rangle_{m'}$$

und

$$(2) \quad X_n * P_{n,m'} \longrightarrow N_S \text{ schwach für } n \longrightarrow \infty$$

für das Standard-Gauss-Mass N_S auf $(T_{\mathcal{E}_{m'}} M, \langle \cdot, \cdot \rangle_{m'})$. $(X_n : n \in \mathbb{N})$ nennen wir Folge der lokal zentralen Zufallsvariablen.

Ist $\langle \cdot, \cdot \rangle_{m'}$ durch eine Bilinearform auf $\mathbb{B} \times \mathbb{B}$ fortsetzbar und bezeichnen wir diese Fortsetzung mit dem gleichen Symbol $\langle \cdot, \cdot \rangle_{m'}$, so gilt folgendes: Wenn Y_n $\mathbb{B}$ -wertige Zufallsvariable sind und

$$X_n - Y_n \longrightarrow 0 \text{ (P}_{n,m'}) \text{ ,}$$

dann ist auch $t \longrightarrow \langle t, Y_n \rangle_{m'}$ ein lokal zentraler Prozess. In diesem Fall werden wir auch $(Y_n : n \in \mathbb{N})$ eine lokal zentrale Folge für die Folge der Experimente F_n nennen. ■

3 Bemerkung

Die Folge der Experimente F_n sei in m' lokal schwach konvergent gegen einen Gauss-Shift für $(T_{\mathcal{E}_{m'}} M, \langle \cdot, \cdot \rangle_{m'})$, und die lokal zentrale Folge sei L_n . Genau dann gilt für eine Folge stochastischer Prozesse $(G_n(t) : t \in T_{\mathcal{E}_{m'}} M)$

$$\frac{dP_{n,m_n}}{dP_{n,m'}} = \exp \left[G_n(t) - \frac{1}{2} |t|_{m'}^2 + o_{P_{n,m'}} \langle n^0 \rangle \right]$$

für $n^{1/2} (m_n - m') \longrightarrow t$, wenn

$$(2) \quad G_n(t) - L_n(t) = o_{P_{n,m'}} \langle n^0 \rangle$$

für alle $t \in T_{\mathcal{E}_{m'}} M$. In diesem Fall gilt auch

$$(3) \quad G_n(t) * P_{n,m'} \longrightarrow N(0, |t|_{m'}^2) \quad \text{schwach für } n \longrightarrow \infty$$

für alle $t \in T_{\mathcal{E}_{m'}} M$.

Dies ergibt sich leicht unter Anwendung des Satzes von Slutsky (vgl. Serfling 15.4). ■

4 Bemerkung

M bezeichne eine differenzierbare Mannigfaltigkeit und

$$F_n := \left[\Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right]$$

seien statistische Experimente. Für $m' \in M$ sei (U, Φ, V) eine Karte von M mit $m' \in U$. Setze $v' := \Phi(m')$ und definiere Experimente

$$\hat{F}_n := \left[\Omega_n, \mathcal{A}_n, \left\{ Q_{n,v} : v \in V_n \right\} \right]$$

mit

$$V_n := \left\{ v \in \mathbb{H} : v' + v n^{-1/2} \in V \right\}$$

und

$$Q_{n,v} := P_{n, \Phi^{-1}(v' + v n^{-1/2})}.$$

Genau dann ist $(F_n : n \in \mathbb{N})$ im m' lokal schwach konvergent gegen einen Gauss-Shift für $(Tg_m M, \langle \cdot, \cdot \rangle_m)$, wenn für jede Folge $(v_n : n \in \mathbb{N}) \subset \mathbb{H}$ mit $\lim v_n = v \in \mathbb{H}$ gilt

$$\frac{dQ_{n,v}}{dQ_{n,0}} = \exp \left[G_n(v) - \frac{1}{2} |v|_{v'}^2 + o_{P_{n,0}}(n^0) \right]$$

mit einer Folge linearer Prozesse $(G_n(v) : v \in \mathbb{H})$, die

$$(2) \quad G_n(v) * Q_{n,0} \longrightarrow N(0, |v|_{v'}^2) \quad \text{schwach für } n \longrightarrow \infty$$

für alle $v \in \mathbb{H}$ erfüllt. $|\cdot|_{v'}$ bezeichnet dabei eine Norm auf $\mathbb{H}$, die von einem von v' abhängigen Skalarprodukt herrührt. Ist eine der beiden äquivalenten Aussagen wahr, so besteht für die lokal zentrale Folge $(L_n : n \in \mathbb{N})$ von $(F_n : n \in \mathbb{N})$ die Beziehung

$$(3) \quad L_n(t) = G_n \left[D\Phi_{m'}(t) \right]$$

und es gilt

$$(4) \quad \langle \cdot, \cdot \rangle_{m'} = \langle D\Phi_{m'}(\cdot), D\Phi_{m'}(\cdot) \rangle_{v'}.$$

Beweis

Wir nehmen an, dass $(F_n : n \in \mathbb{N})$ im m' lokal schwach konvergent gegen einen Gauss-Shift für $(Tg_m M, \langle \cdot, \cdot \rangle_m)$ ist mit lokal zentraler Folge L_n . Definiere für v_n , v und v' wie oben

$$m_n := \tilde{\Phi}^{-1}(v' + v_n n^{-1/2}).$$

Dann gilt

$$n^{1/2} \langle m_n - m' \rangle \longrightarrow (D\tilde{\Phi}_{m'})^{-1}(v).$$

Also ist

$$\frac{dQ_{n,v}}{dQ_{n,o}} = \frac{dP_{n,m}}{dP_{n,m'}} =$$

$$\exp \left[L_n \langle (D\tilde{\Phi}_{m'})^{-1}(v) \rangle - \frac{1}{2} \left| (D\tilde{\Phi}_{m'})^{-1}(v) \right|_{m'}^2 + o_{P_{n,o}}(n^0) \right].$$

Weil $P_{n,m'} = Q_{n,o}$, folgen mit den Definitionen

$$G_n(v) := L_n \langle (D\tilde{\Phi}_{m'})^{-1}(v) \rangle$$

und

$$\langle \cdot, \cdot \rangle_{v'} := \langle (D\tilde{\Phi}_{m'})^{-1}(\cdot), (D\tilde{\Phi}_{m'})^{-1}(\cdot) \rangle_{m'}$$

die Aussagen (1) bis (4).

Umgekehrt sollen jetzt (1) und (2) gelten für irgendeine Karte $(U, \tilde{\Phi}, V)$ mit $m' \in U$. Ist $\langle m_n : n \in \mathbb{N} \rangle$ eine Folge aus M derart, dass

$$n^{1/2} \langle m_n - m' \rangle \longrightarrow t \in T_{\mathbb{E}_{m'}} M,$$

so definiere

$$v_n := n^{1/2} \langle \tilde{\Phi}(m_n) - \tilde{\Phi}(m') \rangle.$$

Dann ist

$$v_n \longrightarrow D\tilde{\Phi}_{m'}(t)$$

und daher

$$\frac{dP_{n,m}}{dP_{n,m'}} = \frac{dQ_{n,v}}{dQ_{n,o}} =$$

$$\exp \left[G_n \left[D\tilde{\Phi}_{m'}(t) \right] - \frac{1}{2} \left| D\tilde{\Phi}_{m'}(t) \right|_{v'}^2 + o_{P_{n,o}}(n^0) \right].$$

Definiere

$$L_n(t) := G_n(D\tilde{\Phi}_{m'}(t))$$

und

$$\langle \cdot, \cdot \rangle_{m'} := \langle D\tilde{\Phi}_{m'}(\cdot), D\tilde{\Phi}_{m'}(\cdot) \rangle_{v'}.$$

Dann folgt, dass $\langle F_n : n \in \mathbb{N} \rangle$ in m' lokal schwach gegen einen Gauss-Shift für $\langle T_{\mathbb{E}_{m'}} M, \langle \cdot, \cdot \rangle_{m'} \rangle$ konvergiert. ■

5 Bemerkung

Aus 4 ergibt sich für den Fall, dass M eine offene Teilmenge eines Hilbert-Raumes H ist:

Ist $(F_n : n \in \mathbb{N})$, $F_n = \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right\}$ in m' lokal

schwach konvergent gegen einen Gauss-Shift für $\left[T_{\xi_{m'}, M}, \langle \cdot, \cdot \rangle_{m'} \right]$, so ist die Folge der in m' lokalisierten Experimente schwach konvergent gegen denselben Gauss-Shift. Umgekehrt ist eine leicht verschärfte Version der schwachen Konvergenz der lokalisierten Experimente hinreichend für die lokal schwache Konvergenz. Genauer gilt: Ist für $m' \in M$

$$\frac{dP_{n, m' + v_n}^{-1/2}}{dP_{n, m'}} = \exp \left[L_n(v) - \frac{1}{2} |v|_{m'}^2 + o_{P_{n, m'}}(n^0) \right]$$

wann immer $v_n \rightarrow v \in H$, und gilt

$$L_n(v) * P_{n, m'} \rightarrow N(0, |t|_0^2) \quad \text{schwach für } n \rightarrow \infty,$$

so konvergiert die Folge $(F_n : n \in \mathbb{N})$ in m' lokal schwach gegen einen Gauss-Shift für $\left[T_{\xi_{m'}, M}, \langle \cdot, \cdot \rangle_{m'} \right]$. ■

6 Lemma

Die Folge der Experimente $F_n = \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right\}$ konvergiere in m' lokal schwach gegen einen Gauss-Shift für $\left[T_{\xi_{m'}, M}, \langle \cdot, \cdot \rangle_{m'} \right]$.

(1) Dann sind die Folgen von Verteilungen

$$\left\{ P_{n, m_n} : n \in \mathbb{N} \right\} \quad \text{und} \quad \left\{ P_{n, m'} : n \in \mathbb{N} \right\} \quad \text{wechselseitig benachbart}$$

für jede Folge $(m_n : n \in \mathbb{N})$ mit $n^{1/2}(m_n - m') \rightarrow t$ für ein $t \in T_{\xi_{m'}, M}$.

(2) Für jede Folge $(m_n : n \in \mathbb{N})$, die

$$n^{1/2}(m_n - m') \rightarrow t_1$$

erfüllt mit einem $t_1 \in T_{\xi_{m'}, M}$, gilt, dass

$$L_n(t_2) * P_{n,m_n} \longrightarrow N(\langle t_1, t_2 \rangle_{m'}, |t_2|_{m'}^2) \text{ schwach}$$

für alle $t_2 \in T_{\xi_{m'}} M$.

Beweis

Die Folge der binären Experimente $\left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,m_n}, P_{n,m'} \right\} \right\}$ konvergiert schwach gegen das binäre Experiment $\left(\Omega, \mathcal{A}, \left\{ P_t, P_o \right\} \right)$, wenn $\left(\Omega, \mathcal{A}, \left\{ P_t : t \in T_{\xi_{m'}} M \right\} \right)$ eine

Darstellung des Gauss-Shifts auf $\left(T_{\xi_{m'}} M, \langle \cdot, \cdot \rangle_{m'} \right)$ ist (vgl. Strasser (1), 60.3). Dieses binäre Grenzexperiment ist wie der Gauss-Shift homogen. Unter Zuhilfenahme des Lemmas 61.3 aus Strasser (1) folgt nun die Aussage.

Wir beweisen nun den Teil (2) des Lemmas. Es ist

$$\left(L_n(t_1), L_n(t_2) \right) * P_{n,m} \longrightarrow N(0, M) \text{ schwach}$$

mit

$$M = \begin{bmatrix} |t_1|_{m'}^2 & \langle t_1, t_2 \rangle_{m'} \\ \langle t_1, t_2 \rangle_{m'} & |t_2|_{m'}^2 \end{bmatrix}$$

(vgl. Strasser (1), 68.3). Also gilt

$$\left(\log \frac{dP_{n,m_n}}{dP_{m'}}, L_n(t_2) \right) * P_{n,m'} \longrightarrow N(a, M) \text{ schwach}$$

mit

$$a = \begin{bmatrix} -\frac{1}{2} |t_1|_{m'}^2 \\ 0 \end{bmatrix}.$$

Daher gilt für jede beschränkte stetige Funktion $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$ mit kompaktem Träger

$$\lim_{n \rightarrow \infty} P_{n,m_n} \left[f \left(\log \frac{dP_{n,m_n}}{dP_{m'}}, L_n(t_2) \right) \right] =$$

$$\lim_{n \rightarrow \infty} P_{n,m'} \left[f \left(\log \frac{dP_{n,m_n}}{dP_{m'}}, L_n(t_2) \right) \frac{dP_{n,m_n}}{dP_{n,m'}} \right] =$$

$$N(a, M) [f(x, y) e^x] =$$

$$N(b, M) [f(x, y)]$$

mit

$$b = \begin{bmatrix} -\frac{1}{2} |t_1|_{m'}^2 \\ \langle t_1, t_2 \rangle_{m'} \end{bmatrix}.$$

Folglich haben wir:

$$\lim_{n \rightarrow \infty} \left[\log \frac{dP_{n, m_n}}{dP_{m'}}, L_n(t_2) \right] * P_{n, m_n} \xrightarrow{1} N(b, M) \quad \text{schwach}$$

und hieraus schliessen wir die Behauptung des Lemmas. ■

7 Lemma

Die Folge der Experimente F_n konvergiere in $m' \in M$ lokal schwach gegen einen Gauss-Shift für $\langle T_{g_{m'} M}, \langle \cdot, \cdot \rangle_{m'} \rangle$. Dann gilt für jede Folge $\langle m_n : n \in \mathbb{N} \rangle$ mit $n^{1/2}(m_n - m') \rightarrow 0$:

$$(1) \quad d_1(P_{n, m_n}, P_{m'}) \rightarrow 0.$$

(2) Für jede Folge kritischer Funktionen ist

$$P_{n, m_n} \varphi_n - P_{m'} \varphi_n \rightarrow 0.$$

(3) Für keine Folge messbarer Abbildungen $f_n : \Omega_n \rightarrow \mathbb{B}$ gilt

$$P_{n, m} \left[n^{1/2} \left(|f_n - m| \right) > \alpha \right] \rightarrow 0$$

lokal gleichmässig um m' für alle $\alpha > 0$.

Beweis

Weil $\left\{ P_{n, m_n} : n \in \mathbb{N} \right\}$ und $\left\{ P_{n, m'} : n \in \mathbb{N} \right\}$ wechselseitig

benachbart sind, folgt für den singulären Anteil aus der Lebesgue-

Zerlegung von P_{n, m_n} bzgl. $P_{n, m'}$

$$(4) \quad P_{n, m_n} \left[N_{m_n} \right] \rightarrow 0$$

(vgl. Strasser (1), 18.9). Ferner ist

$$\frac{dP_{n, m_n}}{dP_{n, m'}} = \exp \left[o_{P_{n, m'}}(n^0) \right].$$

Mit Hilfe des Lemmas Strasser (1), 6.12 und der Beziehung (4) schliessen wir hieraus

$$P_{n,m'} \left[\left| \frac{dP_{n,m_n}}{dP_{n,m'}} - 1 \right| \right] + P_{n,m_n} \left[N_{m_n} \right] \longrightarrow 0.$$

Das ergibt (1), wenn man 11.1 (2) berücksichtigt und weiter folgt dann sogleich (2). Um (3) einzusehen, nimm an, dass

$$P_{n,m} \left[n^{1/2} \left(|f_n - m| \right) > \alpha \right] \longrightarrow 0$$

lokal gleichmässig um m' erfüllt ist. Dann gilt auch

$$P_{n,m_n} \left[n^{1/2} \left(|f_n - m_n| \right) > \alpha \right] \longrightarrow 0.$$

für eine beliebige Folge $(m_n : n \in \mathbb{N})$ mit $n^{1/2}(m_n - m') \longrightarrow t$.

Nun sind $\left\{ P_{n,m_n} : n \in \mathbb{N} \right\}$ und $\left\{ P_{n,m'} : n \in \mathbb{N} \right\}$ wechselseitig benachbart, weswegen aus

$$P_{n,m} \left[n^{1/2} \left(|f_n - m| \right) > \alpha \right] \longrightarrow 0$$

folgt, dass

$$P_{n,m_n} \left[n^{1/2} \left(|f_n - m| \right) > \alpha \right] \longrightarrow 0.$$

Also ist

$$n^{1/2} |f_n - m_n| \longrightarrow 0 \quad P_{n,m_n}\text{-stochastisch einerseits,}$$

andererseits

$$n^{1/2} |f_n - m_n| = n^{1/2} |f_n - m' + m_n - m'| \longrightarrow |t|$$

P_{n,m_n} -stochastisch. Dies ist ein Widerspruch für $t \neq 0$. ■

3.II Die Hauptsätze der Test- und Schätztheorie fuer lokal schwach konvergente Experimentenfolgen

Die folgenden beiden Sätze (als Verallgemeinerungen der beiden Hauptsätze der lokalen asymptotischen Test- und Schätztheorie, vgl. Strasser (2)) machen die Eigenschaft der lokalen schwachen Konvergenz gegen einen Gauss-Shift fuer die statistische Theorie fruchtbar. Zunächst ist aber eine Bezeichnung einzuführen.

8 Definition

Für ein Testproblem (H, K) auf M , $H = M \setminus K$ definiere

$$\begin{aligned}\hat{H}_{m'} &:= \left\{ t \in T_{\mathcal{E}_{m'}} M : t = \lim_{n \rightarrow \infty} n^{1/2} (m_n - m') \text{ für eine Folge} \right. \\ &\quad \left. (m_n : n \in \mathbb{N}) \subset H \right\}. \\ \hat{K}_{m'} &:= T_{\mathcal{E}_{m'}} M \setminus \hat{H}_{m'}.\end{aligned}$$

9 Satz

Gegeben sei eine differenzierbare Hilbert-Mannigfaltigkeit M , ein Testproblem (H, K) auf M und eine Folge statistischer Experimente

$$F_n = \left[\Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right],$$

die in $m' \in M$ lokal schwach gegen einen Gauss-Shift

$$F := \left[\Omega, \mathcal{A}, \left\{ P_t : t \in T_{\mathcal{E}_{m'}} M \right\} \right]$$

für $\left[T_{\mathcal{E}_{m'}} M, \langle \cdot, \cdot \rangle_{m'} \right]$ konvergiert. Dann gilt:

Für jede Folge kritischer Funktionen φ_n für das Testproblem (H, K)

und $F_n = \left[\Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right]$ und jede Teilfolge $\mathbb{N}_1 \subset \mathbb{N}$

existiert ein Test φ für $(\hat{H}_{m'}, \hat{K}_{m'})$ und einen Gauss-Shift F derart, dass für jede Folge $(m_n : n \in \mathbb{N})$ mit

$$n^{1/2} (m_n - m') \longrightarrow t$$

gilt

$$\limsup_{n \in \mathbb{N}_1} P_{n,m_n} \varphi_n \geq P_t \varphi \quad t \in \hat{H}_{m'},$$

$$\liminf_{n \in \mathbb{N}_1} P_{n,m_n} \varphi_n \leq P_t \varphi \quad t \in \hat{K}_{m'}.$$

Beweis

Betrachte eine Folge $(m_n : n \in \mathbb{N})$ wie im Satz. (U, Φ, V) möge eine Karte von M sein mit $m' \in U$. Bezeichne

$$h_n := \Phi(m_n), \quad h' := \Phi(m') \quad \text{und} \quad g := D\Phi_{m'}(t).$$

Dann ist

$$g_n := n^{1/2}(h_n - h') \longrightarrow g$$

und

$$(1) \quad m_n = \Phi^{-1}(h' + n^{-1/2} g_n).$$

Für beliebige kritische Funktionen φ_n wie im Satz gilt

$$(2) \quad \lim_{n \rightarrow \infty} |P_{n, m_n} \varphi_n - P_{n, \Phi^{-1}(h' + n^{-1/2} g)} \varphi_n| = 0.$$

Denn

$$\begin{aligned} & |P_{n, m_n} \varphi_n - P_{n, \Phi^{-1}(h' + n^{-1/2} g)} \varphi_n| \leq \\ & d_1 \left[P_{n, m_n}, P_{n, \Phi^{-1}(h' + n^{-1/2} g)} \right]. \end{aligned}$$

Aufgrund der lokalen schwachen Konvergenz gegen einen Gauss-Shift gilt

$$(3) \quad \lim_{n \rightarrow \infty} \frac{dP_{n, \Phi^{-1}(h' + n^{-1/2} g_n)}}{dP_{n, \Phi^{-1}(h' + n^{-1/2} g)}} = 1 \quad (P_{n, m}).$$

Denn wenn L_n in m' lokal zentral für die Folge der F_n ist, gilt

$$\begin{aligned} & \frac{dP_{n, \Phi^{-1}(h' + n^{-1/2} g_n)}}{dP_{n, m'}} = \\ & \exp \left[L_n \left[(D\Phi_{m'})^{-1}(g) \right] - \frac{1}{2} \left| (D\Phi_{m'})^{-1}(g) \right|_{m'}^2 + o_{P_{n, m'}}(n^0) \right] \end{aligned}$$

infolge von (1) und der lokal schwachen Konvergenz. Ferner ist bei Beachtung der Bemerkung 4

$$\begin{aligned} (4) \quad & \frac{dP_{n, \Phi^{-1}(h' + n^{-1/2} g)}}{dP_{n, m'}} = \\ & \exp \left[L_n \left[(D\Phi_{m'})^{-1}(g) \right] - \frac{1}{2} \left| (D\Phi_{m'})^{-1}(g) \right|_{m'}^2 + o_{P_{n, m'}}(n^0) \right], \end{aligned}$$

da

$$n^{1/2} \left[\Phi^{-1}(h' + n^{-1/2} g) - m' \right] \longrightarrow (D\Phi_{m'})^{-1}(g).$$

Gemäss 6 ist $\left\{ P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)} : n \in \mathbb{N} \right\}$ zu $\left\{ P_{n,m'} : n \in \mathbb{N} \right\}$ benachbart. Deshalb ist auch

$$(5) \quad \lim_{n \rightarrow \infty} \frac{dP_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g_n)}}{dP_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)}} = 1 \quad P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)} \text{-stochastisch}$$

Das Lemma 6.3 aus Strasser (1) liefert nun die gleichgradige Integrierbarkeit der Folge

$$\left\{ \frac{dP_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g_n)}}{dP_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)}} : n \in \mathbb{N} \right\}$$

und dies zusammen mit (5)

$$\lim_{n \rightarrow \infty} P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)} \left[\left| \frac{dP_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g_n)}}{dP_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)}} - 1 \right| \right] = 0.$$

Daraus folgt nun

$$\lim_{n \rightarrow \infty} d_1 \left(P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)}, P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g_n)} \right) = 0,$$

und hieraus schliessen wir (2). Mit Hilfe von (2) ergibt sich jetzt

$$(6) \quad \limsup_{n \in \mathbb{N}_1} P_{n,m_n} \varphi_n = \limsup_{n \in \mathbb{N}_1} P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)} \varphi_n,$$

$$(7) \quad \liminf_{n \in \mathbb{N}_1} P_{n,m_n} \varphi_n = \liminf_{n \in \mathbb{N}_1} P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)} \varphi_n.$$

Die Folge der Experimente

$$\begin{aligned} \bar{F}_n &= \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)} : g \in V_n \right\} \right\} \\ V_n &:= \left\{ g \in H : h' + n^{-1/2}g \in V \right\} \end{aligned}$$

ist schwach konvergent gegen einen Gauss-Shift $\bar{F}$ auf $(H, \langle \cdot, \cdot \rangle)^{-}$ mit

$$\langle \cdot, \cdot \rangle^{-} = \langle (D\tilde{\Phi}_{m'})^{-1} \cdot, (D\tilde{\Phi}_{m'})^{-1} \cdot \rangle_{m'}.$$

Dies folgt sogleich aus 4 oder auch aus (4). Gemäss Strasser (2) (11.8) existiert dann ein Test φ für $\bar{F}$ und $(\bar{H}, \bar{K})$ mit

$$\bar{H} := D\Phi_{m'}(\ddot{H}_{m'}),$$

$$\bar{K} := D\Phi_{m'}(\hat{K}_{m'}),$$

der die folgenden beiden Ungleichungen erfüllt:

$$(8) \quad \limsup_{n \in \mathbb{N}_1} P_{n, \bar{H}^{-1}(h'+n^{-1/2}g)} \varphi_n \geq \bar{P}_g \varphi, \quad g \in \bar{H}$$

und

$$(9) \quad \liminf_{n \in \mathbb{N}_1} P_{n, \bar{H}^{-1}(h'+n^{-1/2}g)} \varphi_n \leq \bar{P}_g \varphi, \quad g \in \bar{K}.$$

Definiere nun

$$(10) \quad P_t := \bar{P}_g, \text{ wenn } g = D\Phi_{m'}(t).$$

Ist dann der Gauss-Shift $\bar{F}$ für $(H, \langle \cdot, \cdot \rangle_{m'})$ gegeben durch

$$\left(\Omega, \mathcal{A}, \left\{ \bar{P}_g : g \in H \right\} \right),$$

so ist

$$\left(\Omega, \mathcal{A}, \left\{ P_t : t \in T_{\mathcal{E}_{m'}, M} \right\} \right)$$

ein Gauss-Shift F für $\left[T_{\mathcal{E}_{m'}, M}, \langle \cdot, \cdot \rangle_{m'} \right]$. Ist $\bar{L}(\cdot)$ zentral für $\bar{F}$,

so ist $\bar{L}(D\Phi_{m'}(\cdot))$ zentral für F . Ist also $(m_n : n \in \mathbb{N})$ wie im Satz, so gilt wegen (6) bis (10)

$$\limsup_{n \in \mathbb{N}_1} P_{n, m_n} \varphi_n \geq \bar{P}_{D\Phi_{m'}(t)} \varphi = P_t \varphi, \quad t \in \hat{H}_{m'}$$

$$\liminf_{n \in \mathbb{N}_1} P_{n, m_n} \varphi_n \leq \bar{P}_{D\Phi_{m'}(t)} \varphi = P_t \varphi, \quad t \in \hat{K}_{m'}. \quad \blacksquare$$

10 Satz

Wir betrachten eine differenzierbare Hilbert-Mannigfaltigkeit M in einem Banach-Raum $\mathbb{B}$ und eine Folge statistischer Experimente

$$F_n = \left(\Omega_n, \mathcal{A}_n, \left\{ P_{n, m} : m \in M \right\} \right),$$

die in $m' \in M$ lokal schwach gegen einen Gauss-Shift

$$F = \left(\Omega, \mathcal{A}, \left\{ P_t : t \in T_{\mathcal{E}_{m'}, M} \right\} \right)$$

für $\left[T_{\mathcal{E}_{m'}, M}, \langle \cdot, \cdot \rangle_{m'} \right]$ konvergiert. $\mathbb{D}$ sei ein weiterer Banach-Raum endlicher Dimension und

$$f : M \longrightarrow \mathbb{D}$$

eine differenzierbare Funktion. Schliesslich sei eine Niveau-kompakte Verlustfunktion

$$W : \mathbb{D} \longrightarrow \mathbb{R}.$$

gegeben. Dann existiert zu jeder Folge von Schätzfunktionen

$$\hat{f}_n : \left[\Omega_n, \mathcal{A}_n \right] \longrightarrow (\mathbb{D}, \mathcal{B}(\mathbb{D}))$$

und jeder Teilfolge $\mathbb{N}_1 \subset \mathbb{N}$ eine randomisierte Schätzfunktion ρ , d.h. ein Markoff-Kern ρ von $(\Omega, \mathcal{A})$ nach $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$, mit

$$(1) \quad \limsup_{n \in \mathbb{N}_1} P_{n, m_n} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right] \geq$$

$$\iint W \left(x - Df_{m'}(t) \right) \rho(\omega, dx) P_t(d\omega)$$

für jede Folge $(m_n : n \in \mathbb{N})$ mit

$$n^{1/2}(m_n - m') \longrightarrow t.$$

Falls

$$\left[n^{1/2} (\hat{f}_n - f(m_n)) \right] * P_{n, m_n} \longrightarrow X_{m'} * P_t \text{ schwach}$$

für eine Zufallsvariable $X_{m'}$ auf $(\Omega, \mathcal{A})$, so ist

$$(2) \quad \limsup_{n \in \mathbb{N}_1} P_{n, m_n} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right] \geq P_t \left[W(X_{m'}) \right].$$

Beweis

Sei (U, Φ, V) eine Karte von M und $m' \in U$. Für alle hinreichend grossen $n \in \mathbb{N}$ sei $h_n := \Phi(m_n)$. Wie im Beweis zu Satz 9 sieht man, dass

$$m_n = \Phi^{-1}(h' + n^{1/2} g_n)$$

für eine Folge $(g_n : n \in \mathbb{N}) \subset \mathbb{H}$ mit

$$\lim_{n \rightarrow \infty} g_n = g = D\Phi_{m'}(t).$$

Also ist

$$P_{n, m_n} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right] =$$

$$P_{n, \Phi^{-1}(h' + n^{-1/2} g_n)} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right].$$

Falls W beschränkt ist, gilt

$$(3) \quad \lim_{n \rightarrow \infty} \left| P_{n, m_n} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right] - P_{n, \bar{\Phi}^{-1}(h' + n^{-1/2} g)} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right] \right| = 0.$$

Denn die Folgenglieder auf der linken Seite sind nicht grösser als

$$|W|_{\infty} P_{n, \bar{\Phi}^{-1}(h' + n^{-1/2} g)} \left[\left| \frac{dP_{n, m_n}}{dP_{n, \bar{\Phi}^{-1}(h' + n^{-1/2} g)}} - 1 \right| \right] + |W|_{\infty} P_{n, m_n} \left[N_{m_n} \right],$$

wobei N_{m_n} die singuläre Menge in der Lebesgue-Zerlegung von P_{n, m_n} bzgl. $P_{n, \bar{\Phi}^{-1}(h' + n^{-1/2} g)}$ bezeichnet. Der letzte Ausdruck ist laut

11.1 (2) gleich

$$2 |W|_{\infty} d_1 \left(P_{n, m_n}, P_{n, \bar{\Phi}^{-1}(h' + n^{-1/2} g)} \right).$$

Wie im Beweis zu Satz 9 sieht man, dass er gegen Null strebt; also gilt (3) und somit

$$(4) \quad \limsup_{n \in \mathbb{N}_1} P_{n, m_n} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right] = \limsup_{n \in \mathbb{N}_1} P_{n, \bar{\Phi}^{-1}(h' + n^{-1/2} g)} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right].$$

Nun ist

$$f(m_n) = f(m') + Df_{m'} \left[(D\bar{\Phi}_{m'})^{-1}(n^{-1/2} g) \right] + o(n^{-1/2}),$$

und daher

$$W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) = W \left(n^{1/2} \left\{ \hat{f}_n - f(m') \right\} + o(n^0) - Df_{m'} \left[(D\bar{\Phi}_{m'})^{-1}(g) \right] \right).$$

Gemäss 4 ist die Folge der Experimente

$$\bar{F}_n = \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n, \bar{\Phi}^{-1}(h' + n^{-1/2} g)} : g \in V_n \right\} \right\},$$

$$V_n := \left\{ g \in \mathbb{H} : h' + n^{-1/2} g \in V \right\},$$

schwach konvergent gegen einen Gauss-Shift $\bar{F}$ für $(\mathbb{H}, \langle \cdot, \cdot \rangle)$ mit

$$(5) \quad \langle \cdot, \cdot \rangle^- = \langle (D\tilde{\Phi}_{m'})^{-1}(\cdot), (D\tilde{\Phi}_{m'})^{-1}(\cdot) \rangle_{m'}.$$

Mit Hilfe der Sätze 62.3, 62.4 und 43.5 aus Strasser (1) folgt nun die Existenz einer randomisierten Schätzfunktion ρ für $\bar{F}$ und das Entscheidungsproblem (H, D, W) mit

$$(6) \quad \limsup_{n \in \mathbb{N}_1} P_{n, m_n} \left[W \left(n^{1/2} (\hat{f}_n - f(m_n)) \right) \right] \geq \\ \iint W \left(x - Df_{m'} \left[(D\tilde{\Phi}_{m'})^{-1}(g) \right] \right) \rho(\omega, dx) \bar{F}_g(d\omega) = \\ \iint W \left(x - Df_{m'} \left[(D\tilde{\Phi}_{m'})^{-1}(g) \right] \right) \rho(\omega, dx) P_t(d\omega),$$

wobei

$$(7) \quad P_t = \bar{P}_{D\tilde{\Phi}_{m'}}(t).$$

Die Gleichung (7) besteht, weil

$$\left\{ \Omega, \mathcal{A}, \left\{ \bar{P}_{D\tilde{\Phi}_{m'}}(t) : t \in T_{\mathcal{E}_{m'}M} \right\} \right\}$$

ein Gauss-Shift auf $\left\{ T_{\mathcal{E}_{m'}M}, \langle \cdot, \cdot \rangle_{m'} \right\}$ ist.

(6) gilt für jede Folge $(m_n : n \in \mathbb{N})$ mit

$$n^{1/2}(m_n - m') \longrightarrow t.$$

Ist nun W unbeschränkt, so ist

$$W = \sup \left\{ V \in C_b(\mathbb{D}) : V \leq W, V \text{ Niveau-kompakt} \right\}.$$

Daraus folgt die Aussage (1) auch für unbeschränkte Verlustfunktionen.

Um die letzte Behauptung des Satzes einzusehen, beachte, dass

$$\left(n^{1/2} (\hat{f}_n - f(m_n)) \right) * P_{n, m_n} = \\ \left(n^{1/2} (\hat{f}_n - f(m')) + o_{P_{n, m'}}(n^0) + Df_{m'}(t) \right) * P_{n, m_n} \longrightarrow \\ X_{m'} * P_t \text{ schwach.}$$

Deshalb ist

$$\left(n^{1/2} (\hat{f}_n - f(m')) + o_{P_{n, m'}}(n^0) \right) * P_{n, \tilde{\Phi}^{-1}(h' + n^{-1/2}g)} \longrightarrow \\ X_{m'} * \bar{P}_g + Df_{m'} \left[(D\tilde{\Phi}_{m'})^{-1}(g) \right] \text{ schwach,}$$

und $X_{m'} + \text{Df}_{m'} \left[(\text{D}\tilde{\Phi}_{m'})^{-1}(\xi) \right]$ ist Häufungspunkt in Verteilung von

$$\left[n^{1/2}(\hat{f}_n - f(m')) + o_{P_{n,m'}}(n^0) \right]. \text{ Mit Hilfe des Satzes 62.4 aus}$$

Strasser (1) und der Gleichung (1) schliessen wir daraus

$$\begin{aligned} \limsup_{n \in \mathbb{N}_1} P_{n,m_n} \left[W \left[n^{1/2}(\hat{f}_n - f(m_n)) \right] \right] = \\ \limsup_{n \in \mathbb{N}_1} P_{n,\tilde{\Phi}^{-1}(h'+n^{-1/2}g)} \left[W \left[n^{1/2}(\hat{f}_n - f(m')) + o_{P_{n,m'}}(n^0) \right. \right. \\ \left. \left. - \text{Df}_{m'} \left[(\text{D}\tilde{\Phi}_{m'})^{-1}(\xi) \right] \right] \right] \geq \end{aligned}$$

$$\bar{P}_g \left[W(X_{m'}) \right] = P_t \left[W(X_{m'}) \right]. \quad \blacksquare$$

3.III Lokale asymptotische Testtheorie

Bevor wir uns nun einzelnen Testproblemen zuwenden, erinnern wir noch an folgenden Sachverhalt (vgl. Strasser (1), 69.15 und 69.14).

H bezeichne einen separablen Hilbert-Raum und $f : H \longrightarrow \mathbb{R}^k$ eine stetige lineare Funktion. Für jeden linearen Prozess $\langle Z(h) : h \in H \rangle$ existiert dann genau eine Zufallsvariable von Ω nach $\mathbb{R}^k$, die wir mit $f \circ Z$ notieren und die

$$x^*(f \circ Z) = Z(f^*(x^*))$$

erfüllt. Dabei ist x^* ein Element aus dem Dualraum $(\mathbb{R}^k)^*$ von $\mathbb{R}^k$ und f^* die zu f duale Abbildung von $\mathbb{R}^{k*}$ nach H^* , dem Dualraum von H . Sie ist gegeben vermöge der Vorschrift

$$f^*(x^*) := x^*(f).$$

Wenn insbesondere $k = 1$, so existiert ein Element $f' \in H$ mit

$$f(h) = \langle f', h \rangle.$$

Dann ist

$$f \circ Z = Z(f').$$

Ist speziell H ein euklidischer Raum, so besitzen die linearen Prozesse auf H eine Darstellung $Z(h) = \langle X, h \rangle$ mit einer Zufallsvariablen X von Ω nach H . In diesem Fall ist

$$f \circ Z = f(X).$$

Mit Hilfe des Satzes 9 können wir für viele Testprobleme Folgen von lokal asymptotisch optimalen Entscheidungsfunktionen angeben. Ein möglicher Optimalitätsbegriff in diesem Zusammenhang ist der der lokalen asymptotischen Zulässigkeit. ■

11 Definition

Sei $\mathcal{F}$ die Menge aller Folgen kritischer Funktionen für die Folge der statistischen Experimente

$$F_n = \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right\}$$

und $\mathcal{F}_0 \subseteq \mathcal{F}$ sowie $m' \in M$. (H, K) sei ein Testproblem für M und die

Mengen $\hat{H}_{m'}$ und $\hat{K}_{m'}$ gemäss 8 gebildet. Eine Folge kritischer

Funktionen $\left\{ \varphi_n : n \in \mathbb{N} \right\}$ heisst in m' lokal asymptotisch zulässig in $\mathcal{F}_0$ für (H, K) , falls für jede Testfolge $\left\{ \psi_n : n \in \mathbb{N} \right\} \in \mathcal{F}_0$ aus

$$(1) \quad \limsup P_{n, m_n} \psi_n \leq \liminf P_{n, m_n} \varphi_n$$

für alle $\langle m_n : n \in \mathbb{N} \rangle \subset M$ mit $n^{1/2}(m_n - m') \longrightarrow t \in \hat{H}_{m'}$ und

$$(2) \quad \liminf P_{n, m_n} \psi_n \geq \limsup P_{n, m_n} \varphi_n$$

für alle $\langle m_n : n \in \mathbb{N} \rangle \subset M$ mit $n^{1/2}(m_n - m') \longrightarrow t \in \hat{K}_{m'}$.

folgt, dass für jede solche Folge von $m_n \in M$

$$(3) \quad \lim P_{n, m_n} (\psi_n - \varphi_n) = 0.$$

Unsere Definition besagt im wesentlichen, dass keine Testfolge aus $\mathcal{F}_0$ eine gleichmässig bessere asymptotische Gütefunktion besitzt als $\left\{ \varphi_n : n \in \mathbb{N} \right\}$, lokal in m' betrachtet. ■

12 Definition

Für eine Folge statistischer Experimente

$$F_n = \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n, m} : m \in M \right\} \right\}$$

und ein Testproblem (H, K) für M sei $\left\{ \varphi_n : n \in \mathbb{N} \right\}$ eine Folge

kritischer Funktionen. Sie heisst in $m' \in M$ lokal asymptotisch unverfälscht zum Niveau α , falls

$$\limsup P_{n, m_n} \varphi_n \leq \alpha$$

für alle $\langle m_n : n \in \mathbb{N} \rangle \subset H$ mit $n^{1/2}(m_n - m') \longrightarrow t \in \hat{H}_{m'}$ und

$$\liminf P_{n, m_n} \varphi_n \geq \alpha$$

für alle $\langle m_n : n \in \mathbb{N} \rangle \subset K$ mit $n^{1/2}(m_n - m') \longrightarrow t \in \hat{K}_{m'}$. Wir nennen die Folge kritischer Funktionen in m' lokal α -ähnlich, falls

$$\lim P_{n, m_n} \varphi_n = \alpha$$

für alle $\langle m_n : n \in \mathbb{N} \rangle \subset M$ mit $n^{1/2}(m_n - m') \longrightarrow t \in \partial \hat{H}_{m'}$.

Dabei bezeichnet $\partial \hat{H}_{m'}$ den topologischen Rand von $\hat{H}_{m'}$. ■

13 Satz

Betrachte eine differenzierbare Hilbert-Mannigfaltigkeit M und eine Folge statistischer Experimente

$$F_n = \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right\}.$$

Für eine stetig differenzierbare Funktion $f : M \rightarrow \mathbb{R}$ werde das Testproblem (H, K) definiert durch

$$H := \{ f(m) \leq 0 \}, \quad K := \{ f(m) > 0 \}.$$

Für ein $m' \in M$ mit $f(m') = 0$ und $Df_{m'} \neq 0$ sei $(F_n : n \in \mathbb{N})$ in m' lokal schwach konvergent gegen einen Gauss-Shift für $(T_{m'} M, \langle \cdot, \cdot \rangle_{m'})$; die lokal zentrale Folge sei L_n . Dann gilt für die Folge kritischer Funktionen

$$\varphi_n := 1_{\{ Df_{m'} \circ L_n > N_{1-\alpha} |Df_{m'}|_{m'} \}},$$

und jede weitere Testfolge $(\psi_n : n \in \mathbb{N})$, die lokal asymptotisch α -ähnlich ist sowie jede Folge $(m_n : n \in \mathbb{N})$ mit

$$n^{1/2} \langle m_n - m' \rangle \rightarrow t$$

für ein $t \in T_{m'} M$, dass

$$(1) \quad \liminf_{n, m_n} P_{n, m_n} \psi_n \geq \lim_{n, m_n} P_{n, m_n} \varphi_n, \quad \text{falls } t \in \hat{H}_{m'},$$

und

$$(2) \quad \limsup_{n, m_n} P_{n, m_n} \psi_n \leq \lim_{n, m_n} P_{n, m_n} \varphi_n, \quad \text{falls } t \in \hat{K}_{m'}.$$

Es ist

$$(3) \quad \left\{ \varphi_n : n \in \mathbb{N} \right\} \text{ lokal asymptotisch unverfälscht vom Niveau } \alpha$$

und auch α -ähnlich. Ausserdem ist

$$(4) \quad \left\{ \varphi_n : n \in \mathbb{N} \right\} \text{ lokal asymptotisch zulässig in der Menge aller Testfolgen, die lokal asymptotisch } \alpha\text{-ähnlich sind.}$$

Beweis

Wir notieren $\|\cdot\|_{m'}$ mit $\|\cdot\|$. Es ist

$$(5) \quad \hat{H}_{m'} = \{ Df_{m'}(t) \leq 0 \}$$

und

$$\hat{K}_{m'} = \{ Df_{m'}(t) > 0 \}.$$

Wenn nämlich $t \in \hat{H}_{m'}$, so gilt wegen $m_n \in H$

$$Df_{m'}(t) = \lim_{n \rightarrow \infty} n^{1/2} Df_{m_n}(m_n - m') \leq$$

$$\limsup_{n \rightarrow \infty} n^{1/2} f(m_n) \leq 0.$$

Umgekehrt erfülle $t \in T_{\mathcal{E}_{m'}} M$ die Bedingung

$$Df_{m'}(t) < 0.$$

Dann haben wir

$$0 > n^{1/2} Df_{m'}(n^{-1/2} t) =$$

$$n^{1/2} f(m' + n^{-1/2} t + o(n^{-1/2})) + o(n^0)$$

und daher $m' + n^{-1/2} t + o(n^{-1/2}) \in H$ für hinreichend grosse

$n \in \mathbb{N}$. Folglich ist $t \in \hat{H}_{m'}$, d.h. $\{ Df_{m'}(t) < 0 \} \subset \hat{H}_{m'}$. Nun ist aber $\{ Df_{m'}(t) = 0 \}$ der Tangentialraum der differenzierbaren

Untermannigfaltigkeit $\{ f(m) = 0 \}$ in m' (vgl. 10.10). Gemäss 10.8

ist dieser in $\hat{H}_{m'}$ enthalten. Daher ist auch $\{ Df_{m'}(t) \leq 0 \} \subset \hat{H}_{m'}$.

Überdies gilt

$$\{ Df_{m'}(t) = 0 \} = \{ t = \lim_{n \rightarrow \infty} n^{1/2}(m_n - m') : f(m_n) = 0 \}.$$

Für das Testen der linearen Funktion $\{ Df_{m'}(t) \leq 0 \}$ gegen

$\{ Df_{m'}(t) > 0 \}$ und einen Gauss-Shift für $\left[T_{\mathcal{E}_{m'}} M, \langle \cdot, \cdot \rangle_{m'} \right]$ ist

der beste ähnliche Test vom Niveau α gegeben durch

$$(6) \quad \varphi = 1_{\{ Df_{m'} \circ L > N_{1-\alpha} |Df_{m'}| \}},$$

wobei L die zentrale Zufallsvariable des Gauss-Shifts bezeichnet

(vgl. Strasser (1), 71.3). Sei nun $\left\{ \psi_n : n \in \mathbb{N} \right\}$ eine weitere

Testfolge für (H, K) und $(F_n : n \in \mathbb{N})$, die in m' lokal asymptotisch

ähnlich zum Niveau α ist. Dann existiert gemäss 9 zu jeder

Teilfolge $N_1 \subset \mathbb{N}$ ein Test ψ für $(\hat{H}_{m'}, \hat{K}_{m'})$ und den Gauss-Shift, so

$$(7) \quad \limsup_{n \in N_1} P_{n, m_n} \psi_n \geq P_t \psi, \quad \text{falls } t \in \hat{H}_{m'},$$

und

$$(8) \quad \liminf_{n \in N_1} P_{n, m_n} \psi_n \leq P_t \psi, \quad \text{falls } t \in \hat{K}_{m'}.$$

Weil $\left\{ \psi_n : n \in \mathbb{N} \right\}$ lokal asymptotisch ähnlich zum Niveau α ist, ist ψ ähnlich zum Niveau α für $(\hat{H}_{m'}, \hat{K}_{m'})$ (vgl. (5)). Gemäss Strasser (1), 71.3 gilt also für den Test (6)

$$(9) \quad P_t \psi \geq P_t \varphi, \text{ wenn } t \in \hat{H}_{m'},$$

und

$$(10) \quad P_t \psi \leq P_t \varphi, \text{ falls } t \in \hat{K}_{m'}.$$

Wähle nun N_1 so, dass

$$\liminf_{n \in N_1} P_{n, m_n} \psi_n = \limsup_{n \in N_1} P_{n, m_n} \psi_n.$$

Dann gilt wegen (7) und (9)

$$\liminf_{n \in N_1} P_{n, m_n} \psi_n \geq P_t \varphi, \text{ wenn } t \in \hat{H}_{m'}.$$

Entsprechend erhalten wir mit einer anderen Teilfolge aufgrund von (8) und (10)

$$\limsup_{n \in N_1} P_{n, m_n} \psi_n \leq P_t \varphi, \text{ wenn } t \in \hat{K}_{m'}.$$

Aus der Definition der Testfolge $\left\{ \varphi_n : n \in \mathbb{N} \right\}$ erhellt sogleich, dass

$$\lim_{n \rightarrow \infty} P_{n, m_n} \varphi_n = P_t \varphi$$

für alle $(m_n : n \in \mathbb{N})$ mit $n^{1/2}(m_n - m') \rightarrow t$. Daraus folgen die Behauptungen (1) und (2) und auch (3). Zum Nachweis von (4) betrachten wir eine Folge kritischer Funktionen ψ_n , die in m' lokal asymptotisch α -ähnlich ist. Wir nehmen für jede Folge $(m_n : n \in \mathbb{N})$ mit $n^{1/2}(m_n - m') \rightarrow t$ die Gültigkeit von

$$(11) \quad \limsup_{n \rightarrow \infty} P_{n, m_n} \psi_n \leq \lim_{n \rightarrow \infty} P_{n, m_n} \varphi_n, \text{ falls } t \in \hat{H}_{m'},$$

und

$$(12) \quad \liminf_{n \rightarrow \infty} P_{n, m_n} \psi_n \geq \lim_{n \rightarrow \infty} P_{n, m_n} \varphi_n, \text{ falls } t \in \hat{K}_{m'},$$

an. Aufgrund von (1) und (2) gelten aber in (11) und (12) auch die umgekehrten Relationen. Also folgt, dass in (11) und (12) bereits die Gleichheit der Grenzwerte gilt und $\lim_{n \rightarrow \infty} P_{n, m_n} \psi_n$ existiert. Hieraus schliessen wir (4). ■

14 Bemerkung

In der Situation des Satzes 13 sei $\hat{m}_n$ eine Folge von

Schätzfunktionen

$$\hat{m}_n : \left[\Omega_n, \mathcal{A}_n \right] \longrightarrow (M, \mathcal{B}(M))$$

derart, dass $n^{1/2}(\hat{m}_n - m')$ lokal zentral für $(F_n : n \in \mathbb{N})$ in m' und schnell konsistent ist. Dann gilt

$$Df_{m'} \circ L = Df_{m'}(n^{1/2}(\hat{m}_n - m')) = n^{1/2}f(\hat{m}_n) + o_{P_{n,m'}}(n^0)$$

und

$$|Df_{m'}|_{m'} = |Df_{\hat{m}_n}|_{\hat{m}_n} + o_{P_{n,m'}}(n^0).$$

Folglich ist

$$\hat{\varphi}_n := 1 \left\{ n^{1/2}f(\hat{m}_n) > N_{1-\alpha} |Df_{\hat{m}_n}|_{\hat{m}_n} \right\}$$

ebenfalls optimal im Sinne des Satzes 13 und unabhängig vom Lokalisationspunkt m' mit $f(m') = 0$, sofern nur $|Df_{\hat{m}_n}|_{\hat{m}_n}$ schwach gegen $|Df_{m'}|_{m'}$ konvergiert (vgl. Abschnitt 5). ■

15 Satz

M sei eine differenzierbare Hilbert-Mannigfaltigkeit und $(F_n : n \in \mathbb{N})$ eine Folge statistischer Experimente $F_n = \left[\Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right]$. Für eine stetig differenzierbare

Funktion $f : M \longrightarrow \mathbb{R}$ werde das Testproblem (H, K) definiert durch

$$H := \{ f(m) = 0 \}, \quad K := \{ f(m) \neq 0 \}.$$

Gegeben sei ein $m' \in H$ mit $\text{rg}(Df_{m'}) = 1$. Die Folge der F_n konvergiere in m' lokal schwach gegen einen Gauss-Shift auf $\left[T_{\xi_{m'}, M}, \langle \cdot, \cdot \rangle_{m'} \right]$. Die lokal zentrale Folge möge L_n sein. Wir definieren eine Folge kritischer Funktionen vermöge

$$\varphi_n := 1 \left\{ |Df_{m'} \circ L_n| > N_{1-\alpha/2} |Df_{m'}|_{m'} \right\}.$$

Sei $(m_n : n \in \mathbb{N})$ eine Folge in M mit

$$n^{1/2}(\hat{m}_n - m') \longrightarrow t.$$

Dann bestehen für jede Folge kritischer Funktionen φ_n , die lokal in m' unverfälscht zum Niveau α ist, die folgenden Beziehungen.

$$(1) \quad \limsup_{n,m_n} P_{n,m_n} \psi_n \leq \lim_{n,m_n} P_{n,m_n} \varphi_n, \quad \text{falls } n^{1/2}(m_n - m') \longrightarrow t.$$

Es ist

$$(2) \quad \left\{ \varphi_n : n \in \mathbb{N} \right\} \quad \text{lokal asymptotisch unverfälscht vom Niveau } \alpha$$

und

$$(3) \quad \left\{ \varphi_n : n \in \mathbb{N} \right\} \quad \text{lokal asymptotisch zulässig in der Menge aller Testfolgen, die lokal asymptotisch unverfälscht zum Niveau } \alpha \text{ sind.}$$

Beweis

Wie in 13, Beweis, setze zur Abkürzung der Notation $\| \cdot \| := \| \cdot \|_{m'}$.

H ist eine differenzierbare Untermannigfaltigkeit von M (vgl. 10.10). Darum ist

$$\hat{H}_{m'} = T_{\mathcal{E}_{m'}} M = \left\{ Df_{m'}(t) = 0 \right\}.$$

Der Test

$$\varphi := 1_{\{ \| Df_{m'} \circ L \| > N_{1-\alpha/2} \| Df_{m'} \| \}}$$

ist gleichmässig optimal unter allen unverfälschten Niveau- α -Tests

für den Gauss-Shift für $\left\{ T_{\mathcal{E}_{m'}} M, \langle \cdot, \cdot \rangle_{m'} \right\}$ und das Testproblem $(\hat{H}_{m'}, \hat{K}_{m'})$ in dem Sinne, dass jeder weitere unverfälschte Niveau- α -Test ψ die Bedingung

$$(4) \quad P_t \psi \leq P_t \varphi$$

erfüllt für alle $t \in T_{\mathcal{E}_{m'}} M$ (vgl. Strasser 71.7). Für die Testfolge

$$\left\{ \varphi_n : n \in \mathbb{N} \right\} \quad \text{gilt nun}$$

$$(5) \quad \lim_{n,m_n} P_{n,m_n} \varphi_n = P_t \varphi,$$

falls $n^{1/2}(m_n - m') \longrightarrow t$. Dies ergibt schon (2). Ist

$\left\{ \psi_n : n \in \mathbb{N} \right\}$ eine weitere in m' lokal asymptotisch unverfälschte Folge kritischer Funktionen, so existiert laut Satz 9 ein Test ψ für den Gauss-Shift und das Testproblem $(\hat{H}_{m'}, \hat{K}_{m'})$ mit

$$(6) \quad \limsup_{n,m_n} P_{n,m_n} \psi_n \geq P_t \psi, \quad \text{falls } t \in \hat{H}_{m'},$$

und

$$(7) \quad \limsup_{n,m_n} P_{n,m_n} \psi_n \leq P_t \psi, \quad \text{falls } t \in \hat{K}_{m'}.$$

Weil $\left[\psi_n : n \in \mathbb{N} \right]$ lokal asymptotisch unverfälscht zum Niveau α ist, ist ψ dies ebenso. Deshalb und wegen (4) gilt

$$(8) \quad \limsup_{n,m_n} P_{n,m_n} \psi_n \leq P_t \varphi, \quad \text{falls } t \in \hat{K}_{m'}.$$

Ausserdem erhalten wir aus der lokalen Unverfälschtheit der Folge

$$\left[\psi_n : n \in \mathbb{N} \right] \text{ und (5)}$$

$$(9) \quad \limsup_{n,m_n} P_{n,m_n} \psi_n \leq P_t \varphi, \quad \text{falls } t \in \hat{H}_{m'}.$$

Hieraus ((8) und (9)) folgt (1). Zum Nachweis von (3) nimm an, dass

$$(10) \quad \limsup_{n,m_n} P_{n,m_n} \psi_n \leq \lim_{n,m_n} P_{n,m_n} \varphi_n, \quad \text{falls } t \in \hat{H}_{m'},$$

und

$$(11) \quad \liminf_{n,m_n} P_{n,m_n} \psi_n \geq \lim_{n,m_n} P_{n,m_n} \varphi_n, \quad \text{falls } t \in \hat{K}_{m'},$$

erfüllt sind. Satz 9 liefert nun die Existenz eines Tests Ψ für den Gauss-Shift und $(\hat{H}_{m'}, \hat{K}_{m'})$ mit

$$(12) \quad \liminf_{n,m_n} P_{n,m_n} \psi_n \geq P_t \Psi, \quad \text{falls } t \in \hat{H}_{m'},$$

und

$$(13) \quad \limsup_{n,m_n} P_{n,m_n} \psi_n \leq P_t \Psi, \quad \text{falls } t \in \hat{K}_{m'}.$$

Aus (9) und (12) gewinnen wir

$$P_t \varphi \geq P_t \Psi, \quad \text{falls } t \in \hat{H}_{m'}$$

und aus (10) und (13)

$$P_t \varphi \leq P_t \Psi, \quad \text{falls } t \in \hat{K}_{m'}.$$

Aber φ ist zulässig in der Menge aller zum Niveau α unverfälschten Tests, und Ψ gehört zu dieser Menge. Folglich gilt

$$(14) \quad P_t \varphi = P_t \Psi, \quad \text{für alle } t \in \text{Tg}_{m'} M.$$

Aus (5), (10), (12) und (14) schliessen wir nun für $t \in \hat{H}_{m'}$ und aus (5), (11), (13) und (14) für $t \in \hat{K}_{m'}$:

$$\lim_{n,m_n} P_{n,m_n} \psi_n = \lim_{n,m_n} P_{n,m_n} \varphi_n \quad \text{für alle } t.$$

Damit ist auch (3) bewiesen. ■

16 Satz

Betrachte eine differenzierbare Hilbert-Mannigfaltigkeit M in einem Banach-Raum und eine Folge statistischer Experimente

$$F_n = \left(\Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right).$$

Gegeben sei das Testproblem

$$H, \quad K := M \setminus H,$$

wobei $H \subset M$ eine differenzierbare Untermannigfaltigkeit sei, deren Tangentialraum $T_{m'}H$ von endlicher Kodimension ist für ein $m' \in H$. Die orthogonale Projektion von $T_{m'}M$ auf $T_{m'}H$ werde mit π bezeichnet. Für m' konvergiere die Folge $(F_n : n \in \mathbb{N})$ lokal schwach gegen einen Gauss-Shift $\left(\Omega, \mathcal{A}, \left\{ P_t : t \in T_{m'}M \right\} \right)$ für $\left(T_{m'}M, \langle \cdot, \cdot \rangle_{m'} \right)$ mit einer zentralen Folge L_n . Definiere eine Testfolge

$$\varphi_n := 1_{\{ (id - \pi) \circ L_n|_{m'} \rangle k_\alpha \}}.$$

Dabei ist k_α bestimmt durch

$$P_0 \left\{ |(id - \pi) \circ L|_{m'} \rangle k_\alpha \right\} = \alpha$$

mit dem zentralen Prozess L des Gauss-Shifts. Sie hat folgende Eigenschaften:

- (1) Die Folge der φ_n ist in m' lokal asymptotisch unverfälscht zum Niveau α .
- (2) Jede weitere in m' lokal asymptotisch unverfälschte Folge kritischer Funktionen ψ_n erfüllt für jedes $c > 0$

$$\inf \left\{ \sup \left\{ \limsup_{n, m_n} P_{n, m_n} \psi_n : n^{1/2}(m_n - m') \longrightarrow t \right\} : \right. \\ \left. |t - \pi(t)|_{m'} = c \right\} \leq \\ \inf \left\{ \inf \left\{ \liminf_{n, m_n} P_{n, m_n} \varphi_n : n^{1/2}(m_n - m') \longrightarrow t \right\} : \right. \\ \left. |t - \pi(t)|_{m'} = c \right\}.$$

- (3) $\left\{ \varphi_n : n \in \mathbb{N} \right\}$ ist in m' lokal asymptotisch zulässig in der Menge der in m' lokal asymptotisch zum Niveau α unverfälschten Testfolgen.

Beweis

Bezeichne kurz $\|\cdot\| := \|\cdot\|_{m'}$. Gemäss 10.9 ist $\hat{H}_{m'} = T_{\mathcal{G}_{m'}} H$. Für das Testproblem $(\hat{H}_{m'}, \hat{K}_{m'})$ und den Gauss-Shift ist der Test

$$\varphi := 1_{\{(\text{Id} - \pi) \circ L\| > k_\alpha\}}$$

unverfälscht zum Niveau α . Wenn $c > 0$, so besteht die Ungleichung

$$(4) \quad \inf \left\{ P_t \psi : \|t - \pi(t)\| = c \right\} \leq \inf \left\{ P_t \varphi : \|t - \pi(t)\| = c \right\}$$

für jeden ebenfalls zum Niveau α unverfälschten Test ψ . Ferner ist dieser Test φ zulässig in der Menge aller zum Niveau α unverfälschten Tests für $(\hat{H}_{m'}, \hat{K}_{m'})$ und einen Gauss-Shift (vgl. Strasser (1), 71.10, 71.12). Es ist

$$(5) \quad \lim_{n, m_n} P_{n, m_n} \varphi_n = P_t \varphi, \quad \text{wenn } n^{1/2}(m_n - m') \longrightarrow t \in T_{\mathcal{G}_{m'}} M.$$

Daher ist (1) wahr. Ist $\left\{ \psi_n : n \in \mathbb{N} \right\}$ wie im Satz in m' lokal asymptotisch unverfälscht zum Niveau α , so existiert gemäss Satz 9 ein Test ψ , so dass für jede Folge von Elementen m_n aus M mit $n^{1/2}(m_n - m') \longrightarrow t \in T_{\mathcal{G}_{m'}} M$ folgendes zutrifft:

$$(6) \quad \limsup_{n, m_n} P_{n, m_n} \psi_n \geq P_t \psi, \quad \text{falls } t \in \hat{H}_{m'},$$

und

$$(7) \quad \limsup_{n, m_n} P_{n, m_n} \psi_n \leq P_t \psi, \quad \text{falls } t \in \hat{K}_{m'}.$$

ψ ist unverfälscht, weil $\left\{ \psi_n : n \in \mathbb{N} \right\}$ lokal asymptotisch unverfälscht zum Niveau α ist. Ausserdem folgt aus (7)

$$(8) \quad \sup \left\{ \limsup_{n, m_n} P_{n, m_n} \psi_n : n^{1/2}(m_n - m') \longrightarrow t \right\} \leq P_t \psi, \quad \text{falls } t \in \hat{K}_{m'}.$$

Nun ist $t \in \hat{K}_{m'}$ gleichbedeutend mit $\|t - \pi(t)\| > 0$. Daraus sowie aus (8), (4) und (5) erhalten wir (2).

Schliesslich nehmen wir an, die Folge $\left\{ \psi_n : n \in \mathbb{N} \right\}$ genüge

$$(9) \quad \limsup_{n, m_n} P_{n, m_n} \psi_n \leq \lim_{n, m_n} P_{n, m_n} \varphi_n, \quad \text{falls } t \in \hat{H}_{m'}, \text{ und}$$

$$(10) \liminf_{n,m} P_{n,m} \psi_n \geq \lim_{n,m} P_{n,m} \varphi_n, \quad \text{falls } t \in \hat{K}_{m'}.$$

Laut Satz 9 gibt es einen Test Ψ für den Gauss-Shift und $(\hat{H}_{m'}, \hat{K}_{m'})$, der

$$(11) \liminf_{n,m} P_{n,m} \psi_n \geq P_t \Psi, \quad \text{falls } t \in \hat{H}_{m'},$$

und

$$(12) \limsup_{n,m} P_{n,m} \psi_n \leq P_t \Psi, \quad \text{falls } t \in \hat{K}_{m'},$$

erfüllt. Aus (9) und (11) folgt unter Berücksichtigung von (5)

$$P_t \varphi \geq P_t \Psi, \quad \text{falls } t \in \hat{H}_{m'},$$

während wir aus (10) und (12)

$$P_t \varphi \leq P_t \Psi, \quad \text{falls } t \in \hat{K}_{m'},$$

erhalten. Ψ ist unverfälscht zum Niveau α , und φ ist zulässig in der Menge aller zum Niveau α unverfälschten Tests. Darum ist

$$(13) P_t \varphi = P_t \Psi, \quad \text{für alle } t \in T_{\mathcal{E}_{m'}} M.$$

Hieraus ergibt sich

$$\lim_{n,m} P_{n,m} \psi_n = \lim_{n,m} P_{n,m} \varphi_n \quad \text{für alle } t$$

und daraus (3). ■

17 Satz

M sei eine differenzierbare euklidische Mannigfaltigkeit und $(F_n : n \in \mathbb{N})$ eine Folge statistischer Experimente

$F_n = \left[\Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right]$. Gegeben sei ein Element $m' \in M$

und die Folge der F_n konvergiere in m' lokal schwach gegen einen

Gauss-Shift $\left[\Omega, \mathcal{A}, \left\{ P_t : t \in T_{\mathcal{E}_{m'}} M \right\} \right]$ auf $\left[T_{\mathcal{E}_{m'}} M, \langle \cdot, \cdot \rangle_{m'} \right]$.

Die lokal zentrale Folge möge L_n sein. Definiere eine Zahl k_α mit Hilfe der zentralen Zufallsvariablen L des Gauss-Shifts durch

$$P_0 \left(\|L\|_{m'}^2 > k_\alpha \right) = \alpha.$$

Das Testproblem sei

$$H = \{m'\}, \quad K = M \setminus \{m'\}.$$

Wir definieren eine Folge kritischer Funktionen durch

$$\varphi_n := 1_{\{ |L_n|_{m'}^2 > k_\alpha \}} \quad .$$

Sei $(m_n : n \in \mathbb{N})$ eine Folge in M mit

$$n^{1/2}(m_n - m') \longrightarrow t.$$

Dann gilt

(1) $\{ \varphi_n : n \in \mathbb{N} \}$ ist lokal asymptotisch unverfälscht vom Niveau α .

(2) Ist $\{ \psi_n : n \in \mathbb{N} \}$ eine weitere lokal asymptotisch unverfälschte Testfolge zum Niveau α , so ist

$$\inf \left\{ \sup \left\{ \limsup P_{n,m_n} \psi_n : n^{1/2}(m_n - m') \longrightarrow t \right\} : \right.$$

$$\left. |t|_{m'} = c \right\} \leq$$

$$\inf \left\{ \inf \left\{ \liminf P_{n,m_n} \varphi_n : n^{1/2}(m_n - m') \longrightarrow t \right\} : \right.$$

$$\left. |t|_{m'} = c \right\},$$

sofern $c > 0$.

(3) $\{ \varphi_n : n \in \mathbb{N} \}$ ist lokal asymptotisch zulässig in der Menge aller Testfolgen, die lokal asymptotisch unverfälscht zum Niveau α sind.

Beweis

Hier ist $\hat{H}_{m'} = \{0\}$ und $\hat{K}_{m'} = T_{\mathcal{E}_{m'}, M} \setminus \{0\}$. Definiere

$|| := ||_{m'}$. Für die zentrale Zufallsvariable L des Gauss-Shifts ist

$$\varphi := 1_{\{ |L|^2 > k_\alpha \}}$$

unverfälscht zum Niveau α für $(\hat{H}_{m'}, \hat{K}_{m'})$, und jeder andere Test ψ , der ebenfalls unverfälscht zum Niveau α für $(\hat{H}_{m'}, \hat{K}_{m'})$ ist, genügt der Ungleichung

$$(4) \quad \inf \left\{ P_t \psi : |t| = c \right\} \leq \inf \left\{ P_t \varphi : |t| = c \right\}$$

für $c > 0$ (vgl. Strasser (1), 30.2 und 30.4). Es ist

$$(5) \quad \lim_{n,m} P_{n,m} \varphi_n = P_t \varphi ,$$

wenn $n^{1/2}(m_n - m') \longrightarrow t \in \text{Tg}_m M$.

Das ergibt (1). Die Aussagen (2) und (3) folgen ganz entsprechend den Aussagen (2) und (3) in 16. ■

3.IV Lokale asymptotische Schätztheorie

Für Schätzprobleme geben wir in den folgenden beiden Sätzen lokal asymptotisch optimale Schätzfolgen an, vorausgesetzt, die Folge der Experimente konvergiert lokal schwach gegen einen Gauss-Shift.

18 Satz

Gegeben sei eine Folge statistischer Experimente

$$F_n = \left[\Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right],$$

wobei M eine differenzierbare Hilbert-Mannigfaltigkeit in einem separablen Banach-Raum ist. Die Folge konvergiere in $m' \in M$ lokal schwach gegen einen Gauss-Shift für $\left[T_{m'} M, \langle \cdot, \cdot \rangle_{m'} \right]$ mit einer lokal zentralen Folge $(L_n : n \in \mathbb{N})$. $f : M \rightarrow \mathbb{R}$ sei eine differenzierbare Funktion. Dann gilt für jede Schätzfolge

$$(1) \quad \hat{f}_n := f(m') + n^{-1/2} Df_{m'} \circ L_n + o_{P_{n,m'}}(n^0) :$$

(2) $\hat{f}_n$ ist in m' lokal asymptotisch mediantreu, d.h. es gilt

$$\liminf P_{n,m_n} \left\{ \hat{f}_n \geq f(m_n) \right\} \geq \frac{1}{2}$$

und

$$\liminf P_{n,m_n} \left\{ \hat{f}_n \leq f(m_n) \right\} \geq \frac{1}{2}$$

für jede Folge $(m_n : n \in \mathbb{N})$ mit $n^{1/2}(m_n - m') \rightarrow t \in T_{m'} M$.

(3) ℓ sei eine nach unten halbstetige beschränkte Funktion, $|\cdot|$ der Betrag auf $\mathbb{R}$ und die Verlustfunktion sei durch $W(\cdot) := \ell \circ |\cdot|$ definiert. Ist dann $(m_n : n \in \mathbb{N})$ eine Folge in M mit $n^{1/2}(m_n - m') \rightarrow t \in T_{m'} M$, so gilt

$$\lim P_{n,m_n} \left[W \left[n^{1/2}(\hat{f}_n - f(m_n)) \right] \right] = N(0, |Df_{m'}|^2)(W).$$

(4) Für eine nach unten halbstetige Verlustfunktion und jede lokal asymptotisch mediantreue Schätzfunktionenfolge $(\tilde{f}_n : n \in \mathbb{N})$ gilt

$$\liminf P_{n,m_n} \left[W \left[n^{1/2}(\tilde{f}_n - f(m_n)) \right] \right] \geq$$

$$\lim P_{n,m_n} \left[W \left[n^{1/2}(\hat{f}_n - f(m_n)) \right] \right],$$

falls $n^{1/2}(m_n - m') \longrightarrow t$.

Beweis

Es gelte

$$n^{1/2}(m_n - m') \longrightarrow t \in T_{\xi_{m'}} M.$$

Dann ist

$$\left[n^{1/2}(\hat{f}_n - f(m')) \right] * P_{n, m_n} = Df_{m'} \circ L_n * P_{n, m_n} \longrightarrow Df_{m'} \circ L * P_t,$$

wobei L den zentralen Prozess des Gauss-Shifts

$\left[\Omega, \mathcal{A}, \left\{ P_t : t \in T_{\xi_{m'}} M \right\} \right]$ bezeichnet. Damit erhalten wir

$$P_{n, m_n} \left\{ \hat{f}_n \geq f(m_n) \right\} =$$

$$P_{n, m_n} \left\{ \left[n^{1/2}(\hat{f}_n - f(m')) \right] \geq \left[n^{1/2}(f(m_n) - f(m')) \right] \right\} =$$

$$P_{n, m_n} \left\{ \left[n^{1/2}(\hat{f}_n - f(m')) \right] + o(n^0) \geq Df_{m'}(t) \right\} \longrightarrow$$

$$P_t \left\{ Df_{m'} \circ L \geq Df_{m'}(t) \right\} = \frac{1}{2}.$$

Ganz entsprechend leitet man die zweite Ungleichung unter (2) her.

Wir haben

$$\left[n^{1/2}(\hat{f}_n - f(m_n)) \right] = Df_{m'} \circ L_n - Df_{m'}(t) + o(n^0)$$

und deshalb

$$\left[n^{1/2}(\hat{f}_n - f(m_n)) \right] * P_{n, m_n} \longrightarrow N(0, |Df_{m'}|_{m'}^2).$$

Dies führt zu der Aussage (3), wenn man beachtet, dass $\mathcal{L}_0[\cdot]$ eine Menge von Unstetigkeitsstellen vom Lebesgue-Mass 0 hat.

Schliesslich sei nun $\tilde{f}_n$ eine beliebige lokal in m' asymptotisch mediantreue Folge von Schätzfunktionen. Gemäss Satz 10 existiert nun eine randomisierte Schätzfunktion ρ für den Gauss-Shift, die

$$\liminf P_{n, m_n} \left[W \left[n^{1/2}(\tilde{f}_n - f(m_n)) \right] \right] \geq$$

$$\iint W \left[x - Df_{m'}(t) \right] \rho(\omega, dx) P_t(d\omega)$$

für $n^{1/2}(m_n - m') \longrightarrow t$ und Niveau-kompakte Verlustfunktionen W

erfüllt. ρ ist mediantreu für $Df_{m'}(t)$. Dann ist gemäss Strasser (1), 72.1, die rechte Seite nicht kleiner als $N(0, \|Df_{m'}\|^2)(W)$. Hieraus folgt unter Berücksichtigung der Beziehung (3) die Behauptung (4) nach Übergang zu beliebigem W wie unter (4). ■

19 Satz

Eine Folge statistischer Experimente

$$F_n = \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,m} : m \in M \right\} \right\}$$

sei gegeben; dabei sei M eine differenzierbare euklidische Mannigfaltigkeit der Dimension $k < \infty$ in einem separablen Banach-Raum B . Die Folge konvergiere in einem $m' \in M$ lokal schwach gegen einen Gauss-Shift für $\left(T_{\bar{m}_n} M, \langle \cdot, \cdot \rangle_{\bar{m}_n} \right)$.

$(L_n : n \in \mathbb{N})$ sei die lokal zentrale Folge und o. E. von der Form $L_n(t) = \langle t, X_n \rangle_{\bar{m}_n}$. Dann gelten für die Schätzfolge

$$(1) \quad \hat{m}_n := m' + n^{-1/2} X_n$$

folgende Aussagen.

$$(2) \quad \lim P_{n,m_n} \left[W \left(n^{1/2} (\hat{m}_n - m_n) \right) \right] = N_{T_{\bar{m}_n} M} [W],$$

wenn $n^{1/2} (m_n - m') \rightarrow t \in T_{\bar{m}_n} M$ und W stetig und beschränkt ist.

(3) Für jede weitere Schätzfolge $(\bar{m}_n : n \in \mathbb{N})$ ist

$$\sup \left\{ \liminf P_{n,m_n} \left[W \left(n^{1/2} (\bar{m}_n - m_n) \right) \right] : \right. \\ \left. n^{1/2} (m_n - m') \text{ konvergiert} \right\} \geq N_{T_{\bar{m}_n} M} [W].$$

Die Verlustfunktion $W : B \rightarrow [0, \infty)$ ist dabei subkonvex und nach unten halbstetig.

(4) Für jede Verlustfunktion W wie in (3), jedes $\delta > 0$ und jede Schätzfolge $(\bar{m}_n : n \in \mathbb{N})$ gilt

$$\liminf_{n \in \mathbb{N}} \sup \left\{ P_{n,m} \left[W \left(n^{1/2} (\bar{m}_n - m) \right) \right] : |m - m'| < \delta \right\} \geq$$

$$N_{T_{\xi_m} M} \left[W \right] .$$

Beweis

Es ist

$$(4) \quad P_{n,m_n} \left[W \left(n^{1/2} (\hat{m}_n - m_n) \right) \right] = \\ P_{n,m_n} \left[W \left(X_n - n^{1/2} (m_n - m') \right) \right] .$$

Nun haben wir

$$(5) \quad \left[X_n - n^{1/2} (m_n - m') \right] * P_{n,m_n} \longrightarrow X * P_0$$

für die zentrale Zufallsvariable X des Gauss-Shifts $\left(\Omega, \mathcal{A}, \left\{ P_t : t \in T_{\xi_m} M \right\} \right)$. Also erhalten wir mit Hilfe von (4) und (5)

$$P_{n,m_n} \left[W \left(n^{1/2} (\hat{m}_n - m_n) \right) \right] \longrightarrow N_{T_{\xi_m} M} \left[W \right] .$$

Die Aussage (2) ist somit wahr.

Laut Satz 10 existiert zu $\bar{m}_n$ eine randomisierte Schätzfunktion ρ mit

$$\liminf_{n,m} P_{n,m} \left[W \left(n^{1/2} (\bar{m}_n - m_n) \right) \right] \geq$$

$$\iint W(x - t) \rho(\omega, dx) P_t(d\omega).$$

Gemäss Strasser (1), 72.5 ist das Supremum über alle $t \in T_{\xi_m} M$ des letzten Ausdrucks nicht kleiner als $N_{T_{\xi_m} M} \left[W \right]$. Dies ergibt die Aussage (3).

Zum Nachweis von (4) geben wir uns irgendein $t \in T_{\xi_m} M$ vor und betrachten eine beliebige Folge $(m_n : n \in \mathbb{N}) \subset M$ mit $n^{1/2}(m_n - m') \longrightarrow t$. Für genügend grosse n ist dann $|m_n - m'| < \delta$ und daher

$$\liminf_{n \in \mathbb{N}} \sup \left\{ P_{n,m} \left[W \left(n^{1/2} (\bar{m}_n - m) \right) \right] : |m - m'| < \delta \right\} \geq$$

$$\liminf_{n \in \mathbb{N}} P_{n,m} \left[W \left(n^{1/2} (\bar{m}_n - m) \right) \right] \geq$$

$$\iint W(x - t) \rho(\omega, dx) P_t(d\omega)$$

gemäss Satz 10. Mithin haben wir auch die Gültigkeit von

$$\liminf_{n \in \mathbb{N}} \sup \left\{ P_{n,m} \left[W \left(n^{1/2} (\bar{m}_n - m) \right) \right] : |m - m'| < \delta \right\} \geq$$

$$\sup \left\{ \iint W(x - t) \rho(\omega, dx) P_t(d\omega) : t \in T_{\mathcal{E}_{m'}, M} \right\}.$$

Der letzte Ausdruck ist laut Strasser (1), 72.5 nicht kleiner als

$$N_{T_{\mathcal{E}_{m'}, M}} \left[W \right].$$

■

4. Lokal schwache Konvergenz im Markoff-Fall

In diesem Abschnitt betrachten wir eine differenzierbare Hilbert-Mannigfaltigkeit M in einem Banachraum B (vgl. 10.6).

$$p(M) := \{ p_m : m \in M \}$$

bezeichne eine Familie von Markoff-Kernen, die alle eine Dichte bzgl. eines Kerns r auf einem polnischen Raum $(E, \mathcal{B}(E))$ besitzen, also $p(M) \subset p(r)$ (vgl. 9.). Die Aussagen 2.4, 2.5 und 2.7 sollen nun auf diesen Fall verallgemeinert werden.

Mit $T_{\mathcal{G}_{m'}} M$ notieren wir den Tangentialraum von M in m' (vgl. 10).

In Verallgemeinerung des Satzes 2.4 erhalten wir:

1 Satz

Sei $m' \in M$ und $P_{1,m'}$ die stationäre Verteilung von $p_{m'}$. Wir nehmen an, die Familie $p(M)$ sei in m' $L^2(P_{2,m'})$ -differenzierbar mit injektiver Ableitung

$$Dp_{m'} : T_{\mathcal{G}_{m'}} M \longrightarrow L^2(P_{2,m'}).$$

Für jedes $t \in T_{\mathcal{G}_{m'}} M$ gelte

$$(1) \quad \frac{1}{n} \sum_{i=1}^n p_{m'} \left[x_i, \left[Dp_{m'}(t)(x_i, \cdot) \right]^2 \right] \longrightarrow P_{2,m'} \left[\left[Dp_{m'}(t) \right]^2 \right] \quad (P_{n,m'}).$$

Dann gilt für jede Familie von Startverteilungen mit

$$(2) \quad \frac{dP_{1,m_n}}{dP_{1,m'}} \longrightarrow 1 \quad (P_{1,m'}) \quad \text{für} \quad |m - m'| \longrightarrow 0$$

und jede Folge $(m_n : n \in \mathbb{N})$ mit

$$(3) \quad n^{1/2}(m_n - m') \longrightarrow t \in T_{\mathcal{G}_{m'}} M:$$

$$(4) \quad \frac{dP_{n,m_n}}{dP_{n,m'}} =$$

$$\exp \left\{ n^{-1/2} \sum_{i=1}^{n-1} Dp_{m'}(t)(x_i, x_{i+1}) - \frac{1}{2} P_{2,m'} \left[\left[Dp_{m'}(t) \right]^2 \right] + \right. \\ \left. \mu_{P_{n,m'}}^0 \right\}.$$

Es ist

$$(5) \quad \left[n^{-1/2} \sum_{i=1}^{n-1} Dp_{m'}(t)(x_i, x_{i+1}) \right] * P_{n,m'} \longrightarrow$$

$$N(0, P_{2,m'} \left[\left(Dp_{m'}(t) \right)^2 \right]) \quad \text{schwach.}$$

Beweis

Sei (U, ψ, V) eine Karte von M und $m' \in U$. Bezeichne $h_n := \psi(m_n)$, $h' := \psi(m')$ und $\phi := \psi^{-1}$. Wir betrachten die Familie

$$q(V) := \left\{ q_h := p_{\phi(h)} : h \in V \right\}$$

von Markoff-Kernen. Setze

$$(6) \quad Q_{n,h} := P_{1,\phi(h)}.$$

Dann ist

$$Q_{n,h} = P_{n,\phi(h)}.$$

Die Familie $q(V)$ ist in $h' \in L^2(Q_{2,h'})$ -differenzierbar mit Ableitung

$$(7) \quad Dq_{h'} : H \longrightarrow L^2(Q_{2,h'}),$$

$$(8) \quad Dq_{h'} = Dp_{m'} \circ (D\psi_{m'})^{-1}$$

gemäss 6.24. Nun gilt

$$\lim n^{1/2}(h_n - h') = D\psi_{m'}(t),$$

also

$$(9) \quad h_n = h' + n^{-1/2} \xi_n$$

mit $\xi_n = n^{1/2}(h_n - h') \longrightarrow D\psi_{m'}(t)$.

Ferner ist

$$(10) \quad \frac{1}{n_i} \sum_{i=1}^n q_{h'} \left[x_i, \left(Dq_{h'}(h)(x_i, \cdot) \right)^2 \right] =$$

$$\frac{1}{n_i} \sum_{i=1}^n p_{m'} \left[x_i, \left(Dp_{m'} \circ (D\psi_{m'})^{-1}(h)(x_i, \cdot) \right)^2 \right] \longrightarrow$$

$$P_{2,m'} \left[\left(Dp_{m'} \left\{ (D\psi_{m'})^{-1}(h) \right\} \right)^2 \right] =$$

$$Q_{2,h'} \left[\left(Dq_{h'}(h) \right)^2 \right].$$

Wegen (6) bis (10) sind die Voraussetzungen des Satzes 2.4 für $q(V)$ erfüllt und wir erhalten

$$\begin{aligned}
(11) \quad & \frac{dP_{n,m}}{dP_{n,m}} = \\
& \frac{dQ_{n,h'} + n^{-1/2} g_n}{dQ_{n,h}} = \\
& \exp \left[n^{-1/2} \sum_{i=1}^{n-1} Dq_{h'} \circ D\psi_{m'}(t) \rangle (x_i, x_{i+1}) - \right. \\
& \quad \left. \frac{1}{2} Q_{2,h'} \left[\left[Dq_{h'} \circ D\psi_{m'}(t) \rangle \right]^2 \right] + o_{Q_{n,h'}}(n^0) \right] = \\
& \exp \left[n^{-1/2} \sum_{i=1}^{n-1} Dp_{m'}(t) \rangle (x_i, x_{i+1}) - \frac{1}{2} P_{2,m'} \left[\left[Dp_{m'}(t) \right]^2 \right] + \right. \\
& \quad \left. o_{P_{n,m'}}(n^0) \right].
\end{aligned}$$

Dies ist gerade die Aussage (4). Wegen (8) ist

$$\begin{aligned}
& \left[n^{-1/2} \sum_{i=1}^{n-1} Dp_{m'}(t) \rangle (x_i, x_{i+1}) \right] \# P_{n,m'} = \\
& \left[n^{-1/2} \sum_{i=1}^{n-1} Dq_{h'} \circ D\psi_{m'}(t) \rangle (x_i, x_{i+1}) \right] \# Q_{n,h'} \xrightarrow{\omega} \\
& N(O, Q_{2,h'}) \left[\left[Dq_{h'} \circ D\psi_{m'}(t) \rangle \right]^2 \right] = \\
& N(O, P_{2,m'}) \left[\left[Dp_{m'}(t) \right]^2 \right].
\end{aligned}$$

Deshalb gilt (5). ■

2 Korollar

Unter den Voraussetzungen des Satzes 1 konvergiert die Folge der Experimente

$$F_n := \left[E^n, \mathcal{B}(E^n), \left\{ P_{n,m} : m \in \mathbb{N} \right\} \right]$$

in m' lokal schwach gegen einen Gauss-Shift auf $\left[T_{\xi_{m'}, M}, \langle \cdot, \cdot \rangle_{m'} \right]$ mit dem Skalarprodukt

$$\langle \cdot, \cdot \rangle_{m'} := P_{2,m'} \left[Dp_{m'}(\cdot) Dp_{m'}(\cdot) \right].$$

Die Folge der Experimente ist also lokal asymptotisch normal.

Beweis

Definition 3.1

■

3 Korollar

Statt der Bedingung (1) in 1 werde die Eindeutigkeit der stationären Verteilung $P_{1,m'}$ von $p_{m'}$ gefordert. Dann sind die Behauptungen von 1 unvermindert wahr.

Beweis

Wie in 2.3 sieht man, dass die Eindeutigkeit der stationären Verteilung $P_{1,m'}$ die Gültigkeit von (1) aus 1 bedingt. ■

Für den Fall zeitlich homogener Markoff-Prozesse erhalten wir

4 Satz

Sei M differenzierbare Hilbert-Mannigfaltigkeit in einem Banach-Raum, $m' \in M$ und $P_{1,m}$ die eindeutig bestimmte stationäre Verteilung für $p_{m'}$. $p(M)$ möge in m' $L^2(P_{2,m'})$ -differenzierbar sein mit injektiver Ableitung

$$Dp_{m'} : T_{\mathfrak{g}_{m'}} M \longrightarrow L^2(P_{2,m'}).$$

Ferner sei

$$(1) \quad m \longrightarrow \frac{dp_m}{dp_{m'}} \text{ stetig an der Stelle } m'$$

und

$$(2) \quad p_{m'}(x, \cdot) \text{ absolut stetig bzgl. } P_{1,m'} \text{ für alle } x \in E.$$

Sind dann $\bar{P}_{1,m}$ beliebige Startverteilungen mit

$$(3) \quad \frac{d\bar{P}_{1,m}}{d\bar{P}_{1,m'}} \longrightarrow 1 \quad (\bar{P}_{1,m'}),$$

so gelten für jede Folge $(m_n : n \in \mathbb{N})$ mit

$$(4) \quad n^{1/2}(m_n - m') \longrightarrow t \in T_{\mathfrak{g}_{m'}} M$$

folgende Beziehungen.

$$(5) \quad \frac{d\bar{P}_{n,m}}{d\bar{P}_{n,m'}}(x_1, \dots, x_n) := \frac{d\bar{P}_{1,m} \otimes_{P_{n-1,m}}}{d\bar{P}_{1,m'} \otimes_{P_{n-1,m'}}}(x_1, \dots, x_n) = \\ \exp \left[n^{-1/2} \sum_{i=1}^{n-1} Dp_{m'}(t)(x_i, x_{i+1}) - \frac{1}{2} P_{2,m'} \left[\left(Dp_{m'}(t) \right)^2 \right] + \right. \\ \left. o_{\bar{P}_{n,m'}}(1) \right].$$

$$(6) \quad \left[n^{-1/2} \sum_{i=1}^{n-1} Dp_{m'}(t)(x_i, x_{i+1}) \right] * \bar{P}_{n,m'} \longrightarrow \\ N(0, P_{2,m'} \left[\left(Dp_{m'}(t) \right)^2 \right]) \quad \text{schwach.}$$

Beweis

Mit Hilfe der Techniken aus dem Beweis zu Satz 1, nämlich dem Übergang zu der Situation des Satzes 2.7 vermittels einer Karte von M , erhält man die Aussagen (5) und (6). ■

5 Korollar

Unter den Voraussetzungen des Satzes 4 konvergiert die Folge der Experimente

$$F_n := \left(E^n, \mathcal{B}(E^n), \left\{ \bar{P}_{n,m} : m \in \mathbb{N} \right\} \right)$$

in m' lokal schwach gegen einen Gauss-Shift auf $\left[Tg_{m'} M, \langle \cdot, \cdot \rangle_{m'} \right]$ mit dem Skalarprodukt

$$\langle \cdot, \cdot \rangle_{m'} := P_{2,m'} \left[Dp_{m'}(\cdot) Dp_{m'}(\cdot) \right].$$

Die Folge der Experimente ist also lokal asymptotisch normal.

5. Anwendungen - Mannigfaltigkeiten als Parameterraume im Markoff-Fall

Diskussion

Die lokale asymptotische Theorie liefert nur Aussagen von lokalem Charakter. Die Folge der lokal zentralen Zufallsvariablen hat viele hervorragende Eigenschaften, hängt aber vom meistens willkürlich gewählten Fusspunkt der Lokalisation der Experimente ab. Von daher liefert die lokale asymptotische Theorie für die Praxis zunächst nur Aussagen über Schranken für optimale statistische Entscheidungsfunktionen, nicht jedoch praktisch verwertbare Schätz- oder Testfolgen selbst. Für Anwendungen wird es darauf ankommen, Tests und Schätzer anzugeben, die vom Lokalisationspunkt unabhängig sind.

Betrachte eine Folge von Experimenten

$$(1) \quad F_n := \left\{ \Omega_n, \mathcal{A}_n, \left\{ P_{n,h} : h \in H_0 \right\} \right\}$$

mit einer offenen Teilmenge H_0 eines euklidischen Raumes. Für $h \in H_0$ sei diese Folge lokal schwach konvergent gegen einen Gauss-Shift. Es gelte

$$\log \frac{dP_{n,h+n^{-1/2}t}}{dP_{n,h}} = \langle t, Z_n \rangle_h - \frac{1}{2} |t|_h^2 + o_{P_{n,h}}(n^0).$$

Das Maximum der Abbildung

$$t \longrightarrow \langle t, Z_n \rangle_h - \frac{1}{2} |t|_h^2$$

wird angenommen in $t = Z_n$. Da nun

$$\max \left\{ \log \frac{dP_{n,g}}{dP_{n,h}} : g \in H_0 \right\} = \max \left\{ \log \frac{dP_{n,h+n^{-1/2}t}}{dP_{n,h}} : t \in H_n \right\}$$

gilt, kann man vermuten, dass

$$(2) \quad n^{1/2} \left[\hat{h}_n - h \right] = Z_n + o_{P_{n,h}}(n^0)$$

besteht für den Maximum-Likelihood-Schätzer $\hat{h}_n$ für h . In Strasser (1), 84. sind Bedingungen formuliert, unter denen für beliebige schwach konvergente Experimentenfolgen wie in (1) die Richtigkeit dieser Vermutung nachgewiesen werden kann. Diese Bedingungen sind

allerdings in praxi nur schwer zu verifizieren. Für den i.i.d.-Fall führt Strasser jedoch vor, wie unter klassischen Annahmen die Beziehung (2) mit Hilfe des oben erwähnten Satzes als zutreffend bewiesen werden kann.

Wir betrachten in diesem Abschnitt folgende Situation:

M sei eine differenzierbare euklidische Mannigfaltigkeit in einem Banach-Raum $\mathbb{B}$ und

$$p(M) := \{ p_m : m \in M \} \subseteq p(r)$$

eine Familie von Markoff-Kernen auf einem polnischen Raum $(E, \mathcal{B}(E))$.

An diese Familie stellen wir folgende Regularitätsbedingungen:

1 Bedingungen 1

(1) Zu jedem p_m , $m \in M$ existiert eine eindeutig bestimmte stationäre Verteilung $P_{1,m}$ und es sei $p_m(x, \cdot)$ absolut stetig bzgl. $P_{1,m}$ für jedes $x \in E$.

(2) Für alle $x \in E$ hängt die Menge $\left\{ y : \frac{dp_m}{dr}(x, y) > 0 \right\}$ nicht von m ab.

(3) Es existiert $D_m \log \frac{dp_m}{dr}(x, y)$ für alle $m \in M$ und alle $x, y \in E$.

2 Bedingungen 2

Sei $\mathcal{U}$ ein Atlas von M . Für jede Karte $(U, \Phi, W) \in \mathcal{U}$ gelten mit den Bezeichnungen $Y := \Phi^{-1}$, $q_v := p_{Y(v)}$ und $Q_{n,v} := P_{n, Y(v)}$, $n \in \mathbb{N}$ folgende Aussagen:

(1) Für $f(x, y, w) := \frac{dq_v}{dr}(x, y)$ existieren alle partiellen Ableitungen 3. Ordnung nach w und sind stetig auf W für alle $x, y \in E$.

(2) Für jedes $w \in W$ existiert eine Umgebung $N(w)$ derart, dass für alle $j, k, l \in \{1, \dots, \dim M\}$ und alle $x \in E$

$$r(x, \sup \left\{ \left| \frac{\partial}{\partial v_j} f(x, y, v) \right| : v \in N(w) \right\}) < \infty,$$

$$r(x, \sup \left\{ \left| \frac{\partial^2}{\partial v_j \partial v_k} f(x, y, v) \right| : v \in N(w) \right\}) < \infty \text{ und}$$

$$Q_{z,v} \left[\sup \left\{ \left| \frac{\partial^3}{\partial v_j \partial v_k \partial v_l} \log f(x,y,v) \right| : v \in N(w) \right\} \right] < \infty.$$

$$(3) \quad Q_{z,v} \left[\left| \frac{\partial}{\partial w_j} \log f(x,y,w) \right|^{z+a} \right] < \infty$$

für ein $a > 0$, das von w abhängen darf.

(4) Die Matrix $\Sigma(w)$ mit den Komponenten

$$\Sigma(w)_{k,l} := Q_{z,v} \left[\frac{\partial}{\partial w_k} \log f(x,y,w) \frac{\partial}{\partial w_l} \log f(x,y,w) \right]$$

ist nicht singulär und ihre Komponenten sind stetig in w . ■

3 Bemerkung

Zwei Atlanten $\mathcal{U}$ und $\mathcal{B}$ von M nennen wir verträglich von der Klasse C^k , falls für beliebige Karten $(U, \Phi, W) \in \mathcal{U}$ und $(U', \Psi, W') \in \mathcal{B}$ mit $U \cap U' \neq \emptyset$ gilt, dass

$$\Psi \Phi^{-1} : \Phi(U \cap U') \longrightarrow W' \text{ und } \Phi \Psi^{-1} : \Psi(U \cap U') \longrightarrow W$$

k -mal stetig differenzierbar sind.

Sind die Bedingungen 2 für einen Atlas $\mathcal{U}$ erfüllt, so auch für jeden weiteren Atlas $\mathcal{B}$, der mit $\mathcal{U}$ verträglich von der Klasse C^3 ist. ■

Bemerkung

Die Bedingungen 1 und 2 sind in Anlehnung an Condition 1.1 und 1.2 in Billingsley (1) formuliert. ■

Für den Rest dieses Abschnitts definieren wir zwei Folgen von Experimenten

$$F_{n,W} := \left[E^n, \mathcal{B}(E^n), \left\{ Q_{n,v} : v \in W \right\} \right]$$

für eine Karte (U, Φ, W) von M und

$$G_{n,M} := \left[E^n, \mathcal{B}(E^n), \left\{ P_{n,m} : m \in M \right\} \right].$$

4 Lemma

Die Bedingungen 1 und 2 seien erfüllt. Dann gilt für eine beliebige Karte (U, Φ, W) aus $\mathcal{U}$:

(1) $F_{n,W}$ ist in jedem $w \in W$ lokal schwach konvergent gegen einen Gauss-Shift für $(\mathbb{R}^k, \langle \cdot, \cdot \rangle_v)$, wobei $k := \dim M$ und

$$\langle \alpha, \beta \rangle_{\mathbf{v}} := \alpha^t \Sigma(\mathbf{w}) \beta.$$

Eine Folge von in $\mathbf{w}$ lokal zentralen Zufallsvariablen ist durch

$$(2) \quad Z_{n,\mathbf{v}} := n^{-1/2} \Sigma(\mathbf{w})^{-1} \sum_{i=1}^{n-1} D_{\mathbf{v}} \log f(x_i, x_{i+1}, \mathbf{w})$$

definiert. Sie erfüllt die Beziehungen

$$(3) \quad Z_{n,\mathbf{v}} * Q_{n,\mathbf{v}} \longrightarrow N(0, \Sigma(\mathbf{w})^{-1})$$

und

$$(4) \quad |Z_{n,\mathbf{v}}|_{\mathbf{v}}^2 * Q_{n,\mathbf{v}} \longrightarrow \chi_k^2,$$

für die zentrale χ^2 -Verteilung mit k Freiheitsgraden.

Ebenso gilt für die Folge der Experimente $G_{n,M}$ und jedes $m \in M$:

(5) $G_{n,M}$ ist in jedem $m \in M$ lokal schwach konvergent gegen einen Gauss-Shift für $\left[T_{\xi_m M}, \langle \cdot, \cdot \rangle_m \right]$. Hier ist

$$\langle s, t \rangle_m := P_{z,m} \left[\left\{ \left[D_m \log \frac{dp_m}{dr} \right] (s) \right\}^t \left[D_m \log \frac{dp_m}{dr} \right] (t) \right].$$

Die Folge der linearen Prozesse

$$(6) \quad L_{n,m}(t) := n^{-1/2} \sum_{i=1}^{n-1} \left[D_m \log \frac{dp_m}{dr} \right] (t), \quad t \in T_{\xi_m M}$$

ist in m lokal zentral.

(7) Ist $m \in U$ für eine Karte $(U, \tilde{\Phi}, \mathbf{w}) \in \mathcal{U}$ und $\tilde{\Phi}(m) := \mathbf{w}$, so gilt für $Z_{n,\mathbf{v}}$ gemäss (2): Die Folge der Zufallsvariablen

$$\hat{Z}_{n,m} := DY_{\mathbf{v}}(Z_{n,\mathbf{v}})$$

ist in m lokal zentral für $G_{n,M}$. Es ist

$$\hat{Z}_{n,m} * P_{n,m} \longrightarrow \left[D\tilde{\Phi}_m \right]^{-1} * N(0, \Sigma(\tilde{\Phi}(m))^{-1}) = N_{T_{\xi_m M}}$$

und

$$|\hat{Z}_{n,m}|_m^2 * P_{n,m} \longrightarrow \chi_k^2.$$

(8) Für die Skalarprodukte bestehen die Relationen

$$\langle \cdot, \cdot \rangle_m = \langle D\tilde{\Phi}_m(\cdot), D\tilde{\Phi}_m(\cdot) \rangle_{\mathbf{v}}$$

und

$$\langle \cdot, \cdot \rangle_{\mathbf{v}} = \langle DY_{\mathbf{v}}(\cdot), DY_{\mathbf{v}}(\cdot) \rangle_m,$$

wenn $\tilde{\Phi}(m) = \mathbf{w}$.

Beweis

Aus den Bedingungen 2 folgt mit Hilfe des Satzes 6.18 die $L^2(Q_{2,v})$ -Differenzierbarkeit der Familie $q(W) := \left\{ q_v : w \in W \right\}$

und ihre $L^2(Q_{2,v})$ -Ableitung in w ist

$$Dq_v(v)(x,y) = v^i D_v \log f(x,y,w) , \quad v \in \mathbb{R}^k.$$

Wegen (4) aus den Bedingungen 2 ist diese Ableitung injektiv. Unter Berücksichtigung der Bedingung 1 (1) erhalten wir hieraus durch Anwendung des Korollars 2.5 die Aussagen (1) und (2). Vermöge 3.1.2 ergeben sich nun (3) und (4).

Mit Hilfe von 6.24 und 6.25 folgt die $L^2(P_{2,m})$ -Differenzierbarkeit der Familie $p(M)$ und zwar mit Ableitung

$$(9) \quad Dp_m(t) = \left[D_v \log f(.,.,w) \right] \circ D\Phi_m(t) = \\ \left[D_m \log \frac{dp_m}{dr} \right] \circ D\Psi_v \circ D\Phi_m(t) = \left[D_m \log \frac{dp_m}{dr} \right] (t).$$

Gemäss 3.1.4 impliziert die lokal schwache Konvergenz der Folge $F_{n,w}$ die lokal schwache Konvergenz der Folge $G_{n,M}$ gegen einen Gauss- Shift auf $(Tg_m M, \langle ., . \rangle_m)$. Dabei ist

$$\langle ., . \rangle_m = \langle D\Phi_m(.), D\Phi_m(.) \rangle_v .$$

Dies ist gerade die Aussage (8). Gemäss (9) ist für $s, t \in Tg_m M$

$$\langle D\Phi_m(s), D\Phi_m(t) \rangle_v = \\ Q_{2,v} \left[\left[D_v \log f(.,.,w) \right] \circ D\Phi_m(s) \left[D_v \log f(.,.,w) \right] \circ D\Phi_m(t) \right] \\ P_{2,m} \left[\left[D_m \log \frac{dp_m}{dr} \right] (s) \left[D_m \log \frac{dp_m}{dr} \right] (t) \right] .$$

Daraus folgt die Behauptung (5). Die Aussage (6) erhalten wir aus (9) und 4.1 bis 4.3. (7) schliessen wir mit Hilfe von 3.1.7 (3) unter Berücksichtigung von

$$L_{n,m}(t) = \langle D\Phi_m(t), Z_{n,v} \rangle_v = \langle t, DY_v \circ Z_{n,v} \rangle_m$$

für $t \in Tg_m M$. ■

5 Lemma

Für $p(M)$ mögen die Bedingungen 1 und 2 erfüllt sein. $\hat{m}_n$ sei eine Lösung der Gleichung

$$\sum_{i=1}^{n-1} D_m \left[\log \frac{dp_m}{dr} \right] (x_i, x_{i+1}) = 0, \quad m \in M$$

und es sei

$$\hat{m}_n - m = o_{P_{n,m}}(n^0)$$

für alle $m \in M$. Dann gilt

$$(1) \quad n^{1/2}(\hat{m}_n - m) \text{ ist in } m \text{ lokal zentral für } G_{n,M}.$$

$$(2) \quad n^{1/2}(\hat{m}_n - m) * P_{n,m} \longrightarrow \left[D\hat{\Phi}_m \right]^{-1} * N(0, \Sigma(w)^{-1}) = N_{T_{\mathcal{E}_m} M}$$

für $w = \hat{\Phi}(m)$.

Ist des weiteren $\hat{w}_n$ eine Folge von Lösungen der Gleichung

$$\sum_{i=1}^{n-1} D_v \left[\log \frac{dq_v}{dr} \right] (x_i, x_{i+1}) = 0, \quad w \in W,$$

die

$$n^{1/2}(\hat{w}_n - w) = o_{Q_{n,v}}(n^0)$$

erfüllt, so sind folgende Aussagen wahr:

$$(3) \quad n^{1/2}(\hat{w}_n - w) \text{ ist in } w \text{ lokal zentral für } F_{n,W}.$$

$$(4) \quad n^{1/2}(\hat{w}_n - w) * Q_{n,v} \longrightarrow N(0, \Sigma(w)^{-1}).$$

$$(5) \quad n^{1/2}(\hat{m}_n - m) - D Y_v \left[n^{1/2}(\hat{w}_n - w) \right] = o_{P_{n,m}}(n^0).$$

Definiere

$$(6) \quad Y_{n,m}(\underline{x}) := \sum_{i=1}^{n-1} D_m \left[\log \frac{dp_m}{dr} \right] (x_i, x_{i+1}).$$

Dann ist

$$(7) \quad 2 \left[Y_{n,\hat{m}_n} - Y_{n,m} \right] * P_{n,m} - |n^{1/2}(\hat{m}_n - m)|_m * P_{n,m} = o_{P_{n,m}}(n^0).$$

Entsprechend gilt für

$$(8) \quad X_{n,v}(\underline{x}) := \sum_{i=1}^{n-1} D_v \left[\log \frac{dq_v}{dr} \right] \langle x_i, x_{i+1} \rangle ,$$

dass

$$(9) \quad 2 \left[X_{n,\hat{v}_n} - X_{n,v} \right] * Q_{n,v} - |n^{1/2}(\hat{w}_n - w)|_v * Q_{n,v} = o_{Q_{n,v}}(n^0) .$$

Beweis

Wir beginnen mit (4). Gemäss Theorem 2.2, Billingsley (1), dessen Voraussetzungen wegen Bedingungen 1 und 2 gegeben sind, gilt (4). Derselbe Satz liefert auch

$$n^{1/2}(\hat{w}_n - w) = \Sigma(w)^{-1} n^{-1/2} \sum_{i=1}^{n-1} D_v \left[\log \frac{dq_v}{dr} \right] \langle x_i, x_{i+1} \rangle = o_{Q_{n,v}}(n^0).$$

Hieraus folgt mit Hilfe des Lemmas 4 (2) und 3.1.2 die Aussage (3). Es ist

$$\begin{aligned} \sum_{i=1}^{n-1} D_v \left[\log \frac{dq_v}{dr} \right] \langle x_i, x_{i+1} \rangle &= 0 \Leftrightarrow \\ \sum_{i=1}^{n-1} D_m \left[\log \frac{dp_m}{dr} \right] \langle x_i, x_{i+1} \rangle \Big|_{m=Y(w)} \circ DY_v &= 0 \Leftrightarrow \\ \sum_{i=1}^{n-1} D_m \left[\log \frac{dp_m}{dr} \right] \langle x_i, x_{i+1} \rangle &= 0, \quad m = Y(w). \end{aligned}$$

Die Voraussetzungen an $\hat{m}_n$ garantieren, dass $\hat{m}_n \in U$ mit gegen 1 strebender $P_{n,m}$ -Wahrscheinlichkeit für jede Karte $(U, \mathfrak{F}, W) \in \mathcal{U}$ mit $m \in U$. Setzt man $\hat{v}_n := \mathfrak{F}(\hat{m}_n)$, so gilt gemäss (3):

$$n^{1/2}(\hat{v}_n - w) * Q_{n,v} \longrightarrow N(0, \Sigma(w)^{-1}) ,$$

also

$$(2) \quad n^{1/2}(\hat{m}_n - m) * P_{n,m} \longrightarrow \left[D\mathfrak{F}_m \right]^{-1} * N(0, \Sigma(w)^{-1}) = N_{T\mathcal{G}_m} M .$$

Die Aussage (1) schliessen wir nun wegen (3) und 4.(7) aus

$$n^{1/2}(\hat{m}_n - m) = DY_v \left[n^{1/2}(\hat{v}_n - w) \right] + o_{P_{n,m}}(n^0).$$

In Billingsley (1), Theorem 2.2 ist die Behauptung (9) bewiesen. Damit folgt auch sogleich (8), weil ja $\hat{m}_n = Y(\hat{v}_n)$ und $\hat{v}_n = \hat{\Phi}(\hat{m}_n)$ gewählt werden kann. ■

6 Satz

$p(M)$ erfülle die Bedingungen 1 und 2. Dann existiert eine Folge von Schätzern $\hat{m}_n : (E^n, \mathcal{B}(E^n)) \longrightarrow M$ mit den Eigenschaften

$$(1) \quad \hat{m}_n - m = o_{P_{n,m}}(n^0)$$

(2) Mit gegen 1 strebender Wahrscheinlichkeit ist $\hat{m}_n$ Lösung der nachstehenden Gleichung in m :

$$(3) \quad \sum_{i=1}^{n-1} D \left[\log \frac{dp_m}{dr} \right] = 0.$$

Falls $\hat{m}_n$ (3) löst, so gilt:

$$(4) \quad n^{1/2}(\hat{m}_n - m) * P_{n,m} \longrightarrow N_{T_{\hat{m}_n} M} \text{ schwach und}$$

(5) $\hat{m}_n$ hat die in 3.IV.19 angeführten Optimalitätseigenschaften.

Beweis

Sei m der wahre Parameter und $(U, \Phi, W) \in \mathcal{U}$ mit $m \in U$. Setze $Y := \Phi^{-1}$. Weil die Bedingungen 1 und 2 erfüllt sind, sind die Conditions 1.1 und 1.2 aus Billingsley (1) erfüllt. Gemäss Theorem 2.1, ebenda, existiert also eine konsistente Schätzfolge $\hat{w}_n$ in W für $\Phi(m) = w$.

Sie löst mit gegen 1 strebender Wahrscheinlichkeit die Gleichung

$$\sum_{i=1}^{n-1} D_v \left[\log \frac{dq_v}{dr} \right] (x_i, x_{i+1}) = 0.$$

Nun ist

$$\sum_{i=1}^{n-1} D_v \left[\log \frac{dq_v}{dr} \right] (x_i, x_{i+1}) = 0 \iff$$

$$\sum_{i=1}^{n-1} D_m \left[\log \frac{dp_m}{dr} \right] (x_i, x_{i+1}) = 0, \quad m \in U.$$

Also erfüllt $\hat{m}_n = Y(w_n)$ die Aussagen (1) und (3).
 Gemäss Lemma 5 gelten (4) und (5). ■

7 Korollar

$p(M)$ erfülle die Bedingungen 1 und 2. $f: M \rightarrow \mathbb{R}$ sei eine differenzierbare Funktion. $\hat{m}_n$ bezeichne die Schätzfolge aus 6. Für sie gilt

(1) $f(\hat{m}_n)$ ist konsistent.

(2) $n^{1/2} [f(\hat{m}_n) - f(m)] * P_{n,m} \rightarrow [Df_m (D\hat{\Phi}_m)^{-1}] * N(0, \Sigma(\hat{\Phi}(m))^{-1})$.

(3) $f(\hat{m}_n)$ ist asymptotisch mediantreu, d.h.

$$\liminf P_{n,m} \{ f(\hat{m}_n) \geq f(m) \} \geq \frac{1}{2}$$

$$\limsup P_{n,m} \{ f(\hat{m}_n) \leq f(m) \} \geq \frac{1}{2}.$$

(4) $f(\hat{m}_n)$ hat die in 3.IV.18 genannten Optimalitätseigenschaften von $\hat{f}_n$ aus 3.IV.18 in jedem $m \in M$. ■

Wir wenden uns nun Testproblemen zu.

8 Satz

$p(M)$ erfülle die Bedingungen 1 und 2. Durch

$$H = \{m\}, K = M \setminus \{m\}$$

sei ein Testproblem definiert. Die Testfolge

$$\varphi_n^* = 1_{\{ |n^{1/2}(\hat{m}_n - m)|_m^2 > \chi_{k,a}^2 \}}$$

mit dem $(1-a)$ -Quantil $\chi_{k,a}^2$ der χ^2 -Verteilung mit $\dim M = k$ Freiheitsgraden und dem konsistenten Maximum-Likelihood-Schätzer $\hat{m}_n$ hat dann folgende Eigenschaften:

(1) $\limsup P_{n,m} \varphi_n = a$.

(2) $\lim P_{n,m'} \varphi_n = 1$ für alle $m' \neq m$.

(3) $(\varphi_n : n \in \mathbb{N})$ hat die in 3.III.17 aufgeführten Optimalitätseigenschaften.

Beweis

Gemäss Lemma 3 (6) ist $G_{n,M}$ in $m \in M$ lokal schwach konvergent gegen einen Gauss-Shift auf $(Tg_m M, \langle \cdot, \cdot \rangle_m)$, und laut Lemma 4 ist $n^{1/2}(\hat{m}_n - m)$ eine Folge lokaler zentraler Zufallsvariablen. Daraus folgt (3). Lemma 4 (7) ergibt (1) und (2). ■

9 Korollar

Die gleichen Aussagen gelten für die Testfolge

$$\varphi_n := 1_{\{Y_{n,\hat{m}_n} - Y_{n,m'} > \chi_{k,\alpha}^2\}}$$

mit der Neyman-Pearson-Statistik wie in Lemma 5 (7). ■

10 Satz

Sei $H \subset M$ eine differenzierbare Untermannigfaltigkeit von M . $p(M)$ und $p(H)$ mögen die Bedingungen 1 und 2 erfüllen. Es sei $\hat{m}_n$ eine konsistente Folge von Maximum-Likelihood-Schätzern für $F_{n,M}$, $\hat{h}_n$ eine konsistente Folge von Maximum-Likelihood-Schätzern für $F_{n,H}$.

Betrachte das Testproblem

$$H, K = M \setminus H$$

und die Testfolge

$$\varphi_n := 1_{\{Y_{n,\hat{m}_n} - Y_{n,\hat{h}_n} > \chi_{l,\alpha}^2\}}$$

mit den $(1-\alpha)$ -Quantil der χ^2 -Verteilung mit $l = \dim M - \dim H$ Freiheitsgraden. Dann gilt:

$$(1) \quad \limsup P_{n,m} \varphi_n \leq \alpha \quad \text{für alle } m \in H.$$

$$(2) \quad \liminf P_{n,m} \varphi_n = 1 \quad \text{für alle } m \in M \setminus H.$$

(3) $(\varphi_n : n \in \mathbb{N})$ hat die in 3.III.16 (1), (2), (3) genannten Optimalitätseigenschaften in jedem $m \in H$.

Beweis

Sei $m \in H$. Dann sind $F_{n,M}$ und $F_{n,H}$ in m lokal schwach konvergent gegen die Gauss-Shifts auf $(Tg_m M, \langle \cdot, \cdot \rangle_m)$ bzw. $(Tg_m H, \langle \cdot, \cdot \rangle'_m)$. Dabei ist $\langle \cdot, \cdot \rangle'_m = \langle \cdot, \cdot \rangle_m \Big|_{Tg_m H \times Tg_m H}$ und wird

deshalb auch kurz mit $\langle \cdot, \cdot \rangle$ bezeichnet.

Bezeichne L_n den in m lokal zentralen Prozess für $F_{n,M}$ wie in 3.III.16 und Π die orthogonale Projektion $T_{\xi_m} M \longrightarrow T_{\xi_m} H$. Gemäss Lemma 5 (8) gilt

$$\begin{aligned} \langle Y_{n,m}^{\wedge} - Y_{n,h_n}^{\wedge} \rangle^{1/2} &= \langle |n^{1/2}(\hat{m}_n - m)|_m \rangle \\ |n^{1/2}(\hat{h}_n - m)|_m &= o_{P_{n,m}}(n^0) \end{aligned}$$

Ferner gilt wegen Lemma 5 (1)

$$|n^{1/2}(\hat{m}_n - m)|_m = |L_n|_m = o_{P_{n,m}}(n^0)$$

und

$$|n^{1/2}(\hat{h}_n - m)|_m = |\Pi L_n|_m = o_{P_{n,m}}(n^0).$$

Also ist

$$Y_{n,m}^{\wedge} - Y_{n,h_n}^{\wedge} = |L_n - \Pi L_n|_m^2 = o_{P_{n,m}}(n^0).$$

Aber $|L_n - \Pi L_n|_m^2$ ist die optimale Teststatistik aus 3.III.16 und

$|L_n - \Pi L_n|_m^2 \longrightarrow \chi_l^2$ schwach. Hieraus folgen die Behauptungen. ■

11 Lemma

Für jede stetig differenzierbare Funktion $f : B \longrightarrow \mathbb{R}$ ist die Abbildung

$$M \ni m \longrightarrow |Df_m|_m$$

stetig bzgl. der Banach-Raum-Norm auf M und des Betrages. $|Df_m|_m$ bezeichnet dabei die Operatornorm des Operators Df_m bzgl. der Norm $|\cdot|_m$ auf $T_{\xi_m} M$.

Beweis

Bezeichnet $w := \tilde{\phi}(m)$ für eine Karte von M , so ist

$$\begin{aligned} (1) \quad |Df_m|_m &= \sup \left\{ \frac{|Df_m(t)|}{|t|_m} : t \in T_{\xi_m} M \right\} = \\ &= \sup \left\{ \frac{|Df_m(DY_v(\xi))|}{|DY_v(\xi)|_m} : \xi \in \mathbb{R}^k \right\} = |D(f \circ Y)_v|_v. \end{aligned}$$

$D(f \circ Y)_v$ ist ein lineares Funktional auf $\mathbb{R}^k$. Folglich existiert ein $\xi \in \mathbb{R}^k$ mit

$$D(f \circ Y)_v = \langle \cdot, \xi \rangle.$$

Ebenso existiert ein $\xi_v \in \mathbb{R}^k$ mit

$$D(f \circ Y)_v = \langle \cdot, \xi_v \rangle_v = \langle \cdot, \Sigma(w) \xi_v \rangle.$$

Daraus folgt

$$\Sigma(w)^{-1} \xi = \xi_v$$

und somit ist

$$(2) \quad |D(f \circ Y)_v|_v^2 = |\xi_v|_v^2 = \xi^t \Sigma(w)^{-1} \xi.$$

Da $w \longrightarrow \Sigma(w)$ stetig ist (vgl. Bedingungen 2), folgt nun die Stetigkeit von

$$w \longrightarrow |D(f \circ Y)_v|_v$$

und hieraus mit Hilfe von (1) die Behauptung. ■

12 Satz

Gegeben sei eine stetig differenzierbare Funktion

$$f: \mathbb{B} \longrightarrow \mathbb{R}.$$

Das Testproblem soll durch

$$H := \{ f(m) \leq 0 \}, \quad K := \{ f(m) > 0 \}$$

definiert sein. Definiere mit dem konsistenten Maximum-Likelihood-Schätzer $\hat{m}_n$ die Testfolge

$$(1) \quad \varphi_n := 1 \left\{ n^{1/2} f(\hat{m}_n) > N_{1-\alpha} \left| Df_{\hat{m}_n} \right|_{\hat{m}_n} \right\}$$

Sie ist in jedem $m' \in M$ mit $f(m') = 0$ lokal asymptotisch optimal im Sinne des Satzes 3.III.13. Darüberhinaus ist sie global konsistent und zulässig in der Menge aller asymptotisch zum Niveau α ähnlichen Testfolgen.

Beweis

Gemäss Satz 3.III.13 ist zu jedem vorgegebenen m' mit $f(m') = 0$ durch

$$(2) \quad \psi_n := 1 \left\{ Df_{m'} \circ L_n > N_{1-\alpha} |Df_{m'}| \right\}$$

eine Folge kritischer Funktionen bestimmt, die die in 3.III.13 angegebenen Optimalitätseigenschaften hat. Dabei sind die

L_n die Glieder einer in m' lokal zentralen Folge. Nun ist

$$(3) \quad n^{1/2} f(\hat{m}_n) = n^{1/2} \left[f(\hat{m}_n) - f(m') \right] =$$

$$Df_{m'}(Y_n) + o_{P_{n,m}}(n^0),$$

wobei ist

$$Y_n := n^{1/2} \left(\hat{m}_n - m' \right)$$

eine lokal zentrale Folge für die Folge der Experimente

$$G_{n,M} := \left\{ E^n, \mathcal{B}(E^n), \left\{ P_{n,m} : m \in M \right\} \right\}$$

ist (vgl. 5 (1)). Weil f stetig differenzierbar ist, gilt gemäss Lemma 11 :

$$(4) \quad \left| Df_{\hat{m}_n} \right|_{\hat{m}_n} \longrightarrow |Df_{m'}| \quad \text{schwach unter } P_{n,m'}.$$

Aus (3) und (4) schliessen wir nun, dass die Testfolgen (1) und (2) asymptotisch identisch verteilt sind. Hieraus ergeben sich die Optimalitätseigenschaften aus Satz 3.III.13 für die Folge der φ_n in m' . Aber m' war beliebig bis auf $f(m') = 0$, und die φ_n hängen von m' nicht ab. Daraus erhalten wir den ersten Teil der Behauptung. Um die Konsistenz der Testfolge nachzuweisen, geben wir uns ein $m \in M$ mit $f(m) > 0$ vor. Dann ist infolge der Konsistenz der Schätzfolge der $\hat{m}_n$

$$n^{1/2} f(\hat{m}_n) \longrightarrow \infty \quad (P_{n,m}),$$

also

$$P_{n,m} \varphi_n \longrightarrow 1 \quad \text{für } m \in K.$$

Analog sieht man

$$P_{n,m} \varphi_n \longrightarrow 0 \quad \text{für } m \in H \setminus \{f(m) = 0\}.$$

Hieraus folgt die Konsistenz der Testfolge. ■

13 Satz

Betrachte eine stetig differenzierbare Funktion

$$f: \mathbb{B} \longrightarrow \mathbb{R}.$$

Das Testproblem soll durch

$$H := \{f(m) = 0\}, \quad K := \{f(m) \neq 0\}$$

definiert sein. Ist die Grösse $\hat{m}_n$ wie in Satz 12 definiert, so ist die Testfolge

$$(1) \quad \varphi_n := 1 \left\{ n^{1/2} \left| f(\hat{m}_n) \right| > N_{1-\alpha/2} \left| Df_{\hat{m}_n} \right|_{\hat{m}_n} \right\}$$

in jedem $m' \in M$ mit $f(m') = 0$ lokal asymptotisch optimal im Sinne des Satzes 3.III.15 und global konsistent. Ferner ist sie asymptotisch zulässig in der Menge aller asymptotisch zum Niveau α unverzerrten Testfolgen.

Beweis

Der Beweis dieser Aussagen erfolgt denen des Satzes 12 entsprechend. ■

14 Beispiel

Die oben gewonnenen Resultate wollen wir nun in einer konkreten Situation anwenden. Unsere Ergebnisse sind teilweise wohlbekannt, vgl. etwa Billingsley (1) oder Basawa/Prakasa Rao. Neu erscheinen in allen Fällen die Aussagen zur Optimalität statistischer Entscheidungsregeln sowie die Behandlung von Schätz- und Testproblemen, die durch differenzierbare Funktionen definiert sind.

Betrachte die Menge

$$M := \left\{ \left(m_{i,j} \right)_{1 \leq i,j \leq N} : m_{i,j} > 0 \text{ für alle } i,j \in \{1, \dots, N\}, \right. \\ \left. \sum_{j=1}^N m_{i,j} = 1 \text{ für alle } i \in \{1, \dots, N\} \right\}$$

für ein $N \in \mathbb{N}$. Jedes Element von M fassen wir als Vektor im Raum $\mathbb{R}^{N^2}$ auf vermöge der Zuordnung $x_{(k-1)N+l} := m_{k,l}$. Diese Identifikation bestehe auch zwischen beliebigen $N \times N$ -Matrizen und Elementen aus $\mathbb{R}^{N^2}$.

M ist eine offene Teilmenge einer $(N^2 - N)$ -dimensionalen linearen Mannigfaltigkeit im $\mathbb{R}^{N^2}$. In jedem Punkt $m \in M$ ist der Tangentialraum $Tg_m M =: T$ der der linearen Mannigfaltigkeit zugrundeliegende lineare Raum.

Es sei $E := \{1, \dots, N\}$, $\mathcal{B}(E)$ die Potenzmenge von E und für jedes $i \in E$ sei $r(i, \cdot)$ das Zählmass auf E . Für $m \in M$ definiere einen Markoff-Kern p_m auf E durch

$$p_m(i, j) := p_m(i, \{j\}) := m_{i,j}.$$

Gemäss Theorem V.2.4, Doob, existieren stationäre Verteilungen für jedes p_m , $m \in M$. Für sie gilt

$$(1) \quad P_{1,m}(i) > 0, \quad i \in \{1, \dots, N\}.$$

Andernfalls wäre für ein i

$$0 = P_{1,m}(i) = \sum_j p_m(i, j) P_{1,m}(j),$$

also $P_{1,m}(j) = 0$ für alle j ; denn $p_m(i, j) > 0$. Aus (1) folgt nun bereits die Eindeutigkeit der stationären Verteilungen $P_{1,m}$ von p_m (vgl. Doob, p. 182, case (g) und case (e)).

Die Bedingungen 1 und 2 sind hier erfüllt und der Maximum-Likelihood-Schätzer ist durch

$$(1) \quad \hat{m}_n := \left[\frac{S_{n-1, k, l}}{S_{n-1, k}}(j_1, \dots, j_n) \right]_{1 \leq k, l \leq N}$$

gegeben. Er hat die Eigenschaften des Satzes 6, ist also insbesondere konsistent, asymptotisch normal und in jedem $m \in M$ lokal asymptotisch optimal. Es gilt

$$(2) \quad n^{1/2} \left(\hat{m}_n - m \right) \otimes P_{n,m} \longrightarrow N_{T,m} \cong N(0, \Pi_T \Sigma(m)^{-1} \Pi_T),$$

wenn $N_{T,m}$ das Standard-Gauss-Mass auf $(T, \langle \cdot, \cdot \rangle_m)$ bezeichnet (vgl.

2). Die Matrix $\Sigma(m)$ ist dabei durch

$$\left[\Sigma(m) \right]_{(i,j),(k,l)} = \frac{P_{i,m} \{i\}}{m_{i,j}} 1_{(i,j)}(k,l).$$

definiert, und Π_T bezeichnet die orthogonale Projektion von $\mathbb{R}^{N^2}$ auf den der linearen Mannigfaltigkeit M zugrundeliegenden linearen Raum. Π_T wird beschrieben durch die Matrix

$$\Pi_T := \langle \Pi_{(i,j),(k,l)} \rangle$$

$$\Pi_{(i,j),(k,l)} := \begin{cases} 0 & , i \neq k \\ -1/N & , i = k, j \neq l \\ 1 - 1/N & , i = k, j = l \end{cases}.$$

Die Resultate aus 6 - 10 und 12 - 13 über Schätz- und Testprobleme sind also auf diesen Fall anwendbar.

6. Differenzierbare Kurven und Familien von Markoff-Kernen

Ist $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ ein statistisches Experiment E , so ist in wichtigen Fällen die Differenzierbarkeit der Familie $\{P_h : h \in H\}$ von Verteilungen P_h eine hinreichende und im wesentlichen auch notwendige Bedingung für die lokale asymptotische Normalität der Folge der Experimente $(\Omega^n, \mathcal{A}^n, \{P_h^n : h \in H\})$ der unabhängigen, identisch verteilten n -fachen Wiederholungen des Experiments E (vgl. Pfanzagl (2), 1.2). Nun soll der Begriff der Differenzierbarkeit von Kurven und Familien von Markoff-Kernen als Verallgemeinerung des entsprechenden Sachverhalts für Verteilungen eingeführt werden.

Wir betrachten in diesem Abschnitt einen polnischen Raum $(E, \mathcal{B}(E))$. Auf ihm sei ein Kern r gegeben (vgl. Bauer, 56). Mit $p(r)$ bezeichnen wir die Menge aller Markoff-Kerne auf $(E, \mathcal{B}(E))$, die eine Dichte bzgl. r besitzen (vgl. 9.1). Eine Kurve $s \rightarrow p_s$ in $p(r)$ ist eine Abbildung von einem offenen Intervall $(-c, c)$ nach $p(r)$. Q sei eine beliebige Verteilung auf $(E, \mathcal{B}(E))$.

1 Definition

Eine Kurve $s \rightarrow p_s$ in $p(r)$ heißt $L^2(Q \otimes r)$ -differenzierbar (an der Stelle 0), falls ein $h \in L^2(Q \otimes r)$ existiert derart, dass

$$Q \otimes r \left[|s|^{-2} \left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_0}{dr} \right)^{1/2} - sh \right\}^2 \right] \rightarrow 0$$

für $s \rightarrow 0$. Die Abbildung $s \rightarrow sh$ nennen wir $L^2(Q \otimes r)$ -Ableitung der Kurve. ■

2 Bemerkung

Die Kurve $s \rightarrow p_s$ ist differenzierbar im Sinne der Definition 1 mit Ableitung $s \rightarrow sh$ dann und nur dann, wenn die Kurve $s \rightarrow Q \otimes p_s$ von Verteilungen auf $(E^2, \mathcal{B}(E^2))$ differenzierbar ist im Sinne der Definition 75.1 aus Strasser (1). ■

3 Satz

Ist $s \rightarrow sh$ die Ableitung einer $L^2(Q \otimes r)$ -differenzierbaren Kurve $s \rightarrow p_s$, so gilt

$$(1) \quad h = g \left(\frac{dp_o}{dr} \right)^{1/2} \quad \text{für ein } g \in L^2(Q\otimes r) \text{ mit}$$

$$(2) \quad p_o(x, g(x, \cdot)) = 0 \text{ Q-f.s.}$$

Beweis

Die erste Behauptung (1) folgt sogleich mit Hilfe des Theorems 75.2 aus Strasser (1) unter Berücksichtigung der Bemerkung 2. Um den zweiten Teil der Aussage nachzuweisen, beachte, dass

$$(3) \quad 0 \leq \lim_{s \rightarrow 0} Q \left(\left| p_o \left[x, s^{-1} \left\{ 2 \left(\frac{dp_s}{dp_o} \right)^{1/2} - 2 - sg \right\} \right] \right| \right) \leq$$

$$\lim_{s \rightarrow 0} Q \otimes p_o \left[\left| s^{-1} \left\{ 2 \left(\frac{dp_s}{dp_o} \right)^{1/2} - 2 - sg \right\} \right| \right] = 0$$

Die Konvergenz des rechten Ausdrucks gegen 0 ergibt sich wie folgt: Wegen (1) und der vorausgesetzten $L^2(Q\otimes r)$ -Differenzierbarkeit ist

$$Q \otimes p_o \left[s^{-2} \left\{ 2 \left(\frac{dp_s}{dp_o} \right)^{1/2} - 2 - sg \right\}^2 \right] \leq$$

$$Q \otimes r \left[s^{-2} \left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - sg \left(\frac{dp_o}{dr} \right)^{1/2} \right\}^2 \right] \rightarrow 0.$$

Da aber $L^2(Q\otimes r)$ -Konvergenz die $L^1(Q\otimes r)$ -Konvergenz impliziert, gilt (3). Somit ist

$$Q(|p_o(x, g(x, \cdot))|) =$$

$$\lim_{s \rightarrow 0} Q \left(\left| p_o \left[x, s^{-1} \left\{ 2 \left(\frac{dp_s}{dp_o} \right)^{1/2} - 2 \right\} \right] \right| \right) \leq$$

$$\lim_{s \rightarrow 0} 2 |s^{-1}| Q(d_2^2(p_s, p_o))$$

gemäss 11.2 (1) für den Hellinger-Abstand der Masse $p_s(x, \cdot)$ und $p_o(x, \cdot)$. Der letzte Ausdruck oben ist wegen 11.2 (5) gleich

$$\lim_{s \rightarrow 0} 2 |s^{-1}| Q \otimes r \left(\left[\left(\frac{dp_s}{dr} \right)^{1/2} - \left(\frac{dp_o}{dr} \right)^{1/2} \right]^2 \right).$$

Aufgrund der $L^2(Q\otimes r)$ -Differenzierbarkeit ist dies

$$\lim_{s \rightarrow 0} |s| \left| Q \otimes r \left(h^2 \right) \right| = 0.$$

Mithin gilt

$$Q \left\{ |p_o(x, g(x, \cdot))| \right\} = 0$$

und hieraus folgt sogleich (2).

Die Eigenschaften (1) und (2) aus obigem Satz beschreiben die Menge der möglichen Ableitungen differenzierbarer Kurven in $p(r)$ vollständig im folgenden Sinne:

4 Satz

Wenn $g \in L^2(Q \otimes p_o)$ und $p_o(x, g(x, \cdot)) = 0$ Q-f.s., dann existiert eine $L^2(Q \otimes r)$ -differenzierbare Kurve $s \rightarrow p_s$ in $p(r)$, deren Ableitung durch $s \rightarrow s g \left(\frac{dp_o}{dr} \right)^{1/2}$ gegeben ist.

Beweis

Wir geben eine Kurve mit Ableitung $s \rightarrow s g \left(\frac{dp_o}{dr} \right)^{1/2}$ an. Definiere

$$(1) \frac{dp_s}{dr}(x, y) :=$$

$$\left\{ \begin{array}{ll} \left[\left(1 - \frac{|s|^2}{4} p_o(x, g^2(x, \cdot)) \right)^{1/2} \left(\frac{dp_o}{dr} \right)^{1/2}(x, y) + \right. \\ \left. \frac{s}{2} g(x, y) \left(\frac{dp_o}{dr} \right)^{1/2}(x, y) \right]^2, & \text{für } 1 - \frac{|s|^2}{4} p_o(x, g^2(x, \cdot)) \geq 0 \\ \frac{dp_o}{dr}(x, y) & , \text{sonst} \end{array} \right.$$

Durch die Definition (1) sind Markoff-Kerne auf $(E, \mathcal{B}(E))$ gegeben. Ferner gilt

$$(2) s^{-1} \left[2 \left(\frac{dp_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - s g \left(\frac{dp_o}{dr} \right)^{1/2} \right] \rightarrow 0 \text{ Q} \otimes r\text{-f.s.}$$

Für solche (x, y) , die $\left(\frac{dp_o}{dr} \right)(x, y) = 0$ erfüllen, ergibt sich (2)

sogleich aus (1). Sei nun $\left(\frac{dp_o}{dr} \right)(x, y) > 0$ und x derart, dass

$p_o(x, g^2(x, \cdot)) < \infty$. Weil $g \in L^2(Q \otimes p_o)$, ist dies für Q-f.a. $x \in E$ erfüllt. Für solche (x, y) existiert ein $s_o := s_o(x, y)$, so dass der erste Fall aus der Definition von $\frac{dp}{dr}^s$ zutrifft, wenn $s \leq s_o$. Es ist dann

$$(3) \quad \left| s^{-1} \left[2 \left(\frac{dp_o}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - s g \left(\frac{dp_o}{dr} \right)^{1/2} \right] \right| = \\ \left| 2 \left(\frac{dp_o}{dr} \right)^{1/2} s^{-1} \left\{ \left[1 - \frac{|s|^2}{4} p_o(x, g^2(x, \cdot)) \right]^{1/2} - 1 \right\} \right| \longrightarrow 0.$$

Also gilt (2). Ausserdem ist

$$\left| s^{-1} \left\{ 2 \left[1 - \frac{|s|^2}{4} p_o(x, g^2(x, \cdot)) \right]^{1/2} \left(\frac{dp_o}{dr} \right)^{1/2} + s g \left(\frac{dp_o}{dr} \right)^{1/2} - \right. \right. \\ \left. \left. 2 \left(\frac{dp_o}{dr} \right)^{1/2} - s g \left(\frac{dp_o}{dr} \right)^{1/2} \right\} \right| \leq \\ 2 \left| \frac{dp_o}{dr} \right|^{1/2} \left| s^{-1} \left\{ \left[1 - \frac{|s|^2}{4} p_o(x, g^2(x, \cdot)) \right]^{1/2} - 1 \right\} \right| \leq \\ 2 \left(\frac{dp_o}{dr} \right)^{1/2} p_o(x, g^2(x, \cdot))^{1/2}.$$

Mithin ist

$$s^{-1} \left| 2 \left(\frac{dp_o}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - s g \left(\frac{dp_o}{dr} \right)^{1/2} \right| \leq \\ \max \left\{ |g| \left(\frac{dp_o}{dr} \right)^{1/2}, 2 \left(\frac{dp_o}{dr} \right)^{1/2} p_o(x, g^2(x, \cdot)) \right\} \in L^2(Q \otimes r).$$

Der Satz von der majorisierten Konvergenz ergibt nun, dass $s \longrightarrow p_s$ differenzierbar mit Ableitung $s \longrightarrow s g \left(\frac{dp_o}{dr} \right)^{1/2}$ ist. ■

Aus den Sätzen 3 und 4 folgt, dass die Menge der möglichen Ableitungen differenzierbarer Kurven $s \longrightarrow p_s$ in $p(r)$ genau die Menge der Abbildungen $s \longrightarrow sh$ mit $h = g \left(\frac{dp_o}{dr} \right)^{1/2}$ und

$p_o(x, g(x, \cdot)) = 0$ Q-f.s. sowie $g \in L^2(Q \otimes p_o)$ ist. Dies gibt Anlass zu folgender Definition.

5 Definition

Die Menge

$$T_{p_o}(r) := \left\{ g \in L^2(Q \otimes p_o) : p_o(x, g(x, \cdot)) = 0 \text{ Q-f.s.} \right\}$$

heisst Tangentialraum von p_o . Die Elemente dieser Menge nennen wir Tangentialvektoren an p_o . Insbesondere sagen wir, eine Kurve

$s \rightarrow p_s$ in $p(r)$ ist $L^2(Q \otimes r)$ -differenzierbar mit Tangentialvektor $g \in T_{p_o}(r)$, falls die Kurve $L^2(Q \otimes r)$ -differenzierbar mit Ableitung

$$s \rightarrow g \left(\frac{dp_o}{dr} \right)^{1/2} \text{ ist.} \quad \blacksquare$$

Wir bemerken noch, dass $T_{p_o}(r)$ ein abgeschlossener linearer Teilraum von $L^2(Q \otimes p_o)$ ist.

Im Satz 4 und dessen Beweis haben wir zu jedem $g \in T_{p_o}(r)$ eine $L^2(Q \otimes r)$ -differenzierbare Kurve mit Tangentialvektor g angegeben. Wenn g beschränkt ist, gibt es besonders einfach zu definierende Kurven mit dieser Eigenschaft.

6 Bemerkung

Ist $g \in L^m(Q \otimes r)$ und $p_o(x, g(x, \cdot)) = 0$ Q-f.s., so ist für $|s| < |g|_\infty^{-1}$ durch

$$(1) \quad \frac{dp_o}{dr} := \begin{cases} \left[1 + sg(x, y) \right] \left(\frac{dp_o}{dr} \right)(x, y) & , p_o(x, g(x, \cdot)) = 0 \text{ und } g(x, y) \leq |g|_\infty \\ \frac{dp_o}{dr}(x, y) & , \text{sonst} \end{cases}$$

eine $L^2(Q \otimes r)$ -differenzierbare Kurve in $p(r)$ gegeben, deren Tangentialvektor g ist.

Beweis

Ohne Einschränkung sei $|g(x,y)| \leq |g|_\infty$ für alle $(x,y) \in E^2$. Wenn nun $s \leq |g|_\infty^{-1}$ ist, so handelt es sich bei den unter (1) definierten Abbildungen um Dichten von Markoff-Kernen. Da ausserdem

$$(2) \quad s^{-1} \left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_0}{dr} \right)^{1/2} - sg \left(\frac{dp_0}{dr} \right)^{1/2} \right\} (x,y) = \\ s^{-1} \left\{ 2 \left(1 + sg(x,y) \right)^{1/2} - 2 - sg(x,y) \right\} \left(\frac{dp_0}{dr} \right)^{1/2} (x,y) \xrightarrow{s \rightarrow 0} 0$$

für alle y und Q -f.a. x und durch $3|g|$ eine Majorante für die Integranden aus (2) gegeben ist, sofern nur s genügend klein ist, ergibt sich die Aussage der Bemerkung unter Zuhilfenahme des Satzes von der majorisierten Konvergenz. ■

Die folgende Aussage betrifft das Verhalten der singulären Anteile aus der Lebesgue-Zerlegung $\left(\frac{dp_s}{dp_0}, \langle N_s(x) : x \in E \rangle \right)$ von p_s

bzgl. p_0 (vgl. 9.5). Es gilt

7 Satz

Ist eine Kurve $s \rightarrow p_s$ in $p(r)$ $L^2(Q(r))$ -differenzierbar, so ist

$$\lim_{s \rightarrow 0} s^{-2} Q \left[p_s(x, N_s(x)) \right] = 0.$$

Beweis

Bezeichne

$$s \rightarrow sg \left(\frac{dp_0}{dr} \right)^{1/2} =: sh$$

die Ableitung von $s \rightarrow p_s$. Weil

$$\left(\frac{dp_0}{dr} \right) (x, \cdot) 1_{N_s(x)}(\cdot) = 0 \quad r(x, \cdot)\text{-f.s.},$$

ist

$$p_s(x, N_s(x)) = r \left[x, \left(\frac{dp_s}{dr} \right) (x, \cdot) 1_{N_s(x)}(\cdot) \right] = \\ \frac{1}{4} r \left[x, \left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} (x, \cdot) - 2 \left(\frac{dp_0}{dr} \right)^{1/2} (x, \cdot) \right\}^2 1_{N_s(x)}(\cdot) \right] \leq$$

$$r \left[x, \left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} (x, \cdot) - 2 \left(\frac{dp_o}{dr} \right)^{1/2} (x, \cdot) - sh(x, \cdot) \right\}^2 1_{N_s(x)}(\cdot) \right] + \\ r \left[x, s^2 h^2(x, \cdot) 1_{N_s(x)}(\cdot) \right].$$

Somit gilt

$$s^{-2} Q[p_s(x, N_s(x))] \leq \\ s^{-2} Q \otimes r \left[\left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - sh \right\}^2 1_{N_s(\cdot)} \right] + \\ s^{-2} Q \otimes r \left[s^2 h^2 1_{N_s(\cdot)} \right].$$

Der erste Summand des letzten Ausdrucks verschwindet für $s \rightarrow 0$ aufgrund der vorausgesetzten Differenzierbarkeit der Kurve. Der zweite Summand ist gleich

$$Q \otimes r \left[s^2 \left(\frac{dp_o}{dr} \right) 1_{N_s(\cdot)} \right] = 0$$

aufgrund der Beziehung (1). Hieraus folgt nun die Behauptung. ■

Der Begriff einer differenzierbaren Kurve in $p(r)$ kann auch etwas anders eingeführt werden. Es stellt sich allerdings alsbald heraus, dass beide Differenzierbarkeitsbegriffe gleichwertig sind.

8 Definition

Eine Kurve $s \rightarrow p_s \in p(r)$ heisst $L^2(Q \otimes p_o)$ -differenzierbar, falls ein $g \in L^2(Q \otimes p_o)$ existiert, so dass

$$(1) \lim_{s \rightarrow 0} s^{-2} Q \otimes p_o \left[\left\{ 2 \left(\frac{dp_s}{dp_o} \right)^{1/2} - 2 - sg \right\}^2 \right] = 0$$

und

$$(2) \lim_{s \rightarrow 0} s^{-2} Q \left[p_s(x, N_s(x)) \right] = 0$$

für die Lebesgue-Zerlegungen $\left(\frac{dp_s}{dp_o}, (N_s(x) : x \in E) \right)$ von p_s

bzgl. p_o . Die Abbildung $s \rightarrow sg$ heisst die $L^2(Q \otimes p_o)$ -Ableitung von $s \rightarrow p_s$, g heisst Tangentialvektor der Kurve. ■

9 Satz

Folgende Aussagen sind äquivalent:

(1) $s \longrightarrow p_s$ ist $L^2(Q\otimes r)$ -differenzierbar mit Ableitung

$$s \longrightarrow sh = sg \left[\frac{dp_o}{dr} \right]^{1/2}.$$

(2) $s \longrightarrow p_s$ ist $L^2(Q\otimes p_o)$ -differenzierbar mit Ableitung $s \longrightarrow sg$.

Beweis

Beide Aussagen folgen, wenn man sich darauf besinnt, dass die Menge $\{y : \left[\frac{dp_o}{dr} \right](x,y) = 0\}$ für jedes x als singuläre Menge der Lebesgue-Zerlegung dienen kann und dabei 9.7 im Auge behält. Dann kann man die Grenzwerte wechselweise gegeneinander abschätzen. ■

10 Korollar

Der Begriff der $L^2(Q\otimes r)$ -Differenzierbarkeit einer Kurve $s \longrightarrow p_s$ ist unabhängig von dem dominierenden Kern r . ■

Das folgende Lemma gibt Auskunft in der Frage, wie sich die Abstände zweier Kurven in $p(r)$ verhalten müssen, um aus der $L^2(Q\otimes p_o)$ -Differenzierbarkeit der einen Kurve auf die gleiche Eigenschaft der anderen Kurve zu schliessen.

11 Lemma

Ist $s \longrightarrow p_s$ eine $L^2(Q\otimes p_o)$ -differenzierbare Kurve in $p(r)$ mit Tangentialvektor g , so ist eine zweite Kurve $s \longrightarrow q_s$ in $p(r)$ mit $q_o = p_o$ $L^2(Q\otimes p_o)$ -differenzierbar mit demselben Tangentialvektor g dann und nur dann, wenn für die Affinitäten der Masse $Q\otimes p_s$ und $Q\otimes q_s$ gilt

$$(1) \alpha(Q\otimes p_s, Q\otimes q_s)^{(s^{-2})} \longrightarrow 1 \text{ für } s \longrightarrow 0.$$

Beweis

Wir nehmen an, dass beide Kurven differenzierbar mit demselben Tangentialvektor sind. Dann ist für $s \longrightarrow 0$

$$(2) \quad s^{-2} d_2^2 \langle Q \otimes p_s, Q \otimes q_s \rangle \leq$$

$$\begin{aligned} & \frac{s^{-2}}{4} Q \otimes r \left[\left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - s \xi \left(\frac{dp_o}{dr} \right)^{1/2} \right\}^2 \right] + \\ & \frac{s^{-2}}{4} Q \otimes r \left[\left\{ 2 \left(\frac{dq_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - s \xi \left(\frac{dp_o}{dr} \right)^{1/2} \right\}^2 \right] \longrightarrow 0. \end{aligned}$$

Bezeichnet nun

$$\alpha_s := \alpha \langle Q \otimes q_s, Q \otimes p_s \rangle,$$

so ist

$$\begin{aligned} \alpha_s^{(s^{-2})} &= \exp \left[s^{-2} \log \alpha_s \right] = \\ & \exp \left[s^{-2} \left[(1 - \alpha_s) - \frac{1}{2} (1 - \alpha_s)^2 + \dots \right] \right] = \\ & \exp \left[s^{-2} d_2^2 \langle Q \otimes p_s, Q \otimes q_s \rangle \left[\sum_{k=1}^{\infty} \frac{1}{k!} (-1)^{k+1} (1 - \alpha_s)^{k-1} \right] \right]. \end{aligned}$$

Die Reihe in dem letzten Ausdruck konvergiert gleichmässig in s . Dies zusammen mit (2) führt zu (1).

Die umgekehrte Implikation beweisen wir folgendermassen. Ist $s \longrightarrow p_s$ $L^2 \langle Q \otimes p_o \rangle$ -differenzierbar mit Tangentialvektor ξ und ist (1) erfüllt, so gilt

$$\begin{aligned} 0 &\leq 1 - \exp \left[- s^{-2} d_2^2 \langle Q \otimes p_s, Q \otimes q_s \rangle \right] = \\ & 1 - \left[\exp(\alpha_s - 1) \right]^{(s^{-2})} \leq 1 - \alpha_s^{(s^{-2})} \longrightarrow 0 \text{ für } s \longrightarrow 0. \end{aligned}$$

Hieraus erhalten wir

$$(3) \quad s^{-2} d_2^2 \langle Q \otimes p_s, Q \otimes q_s \rangle \longrightarrow 0, \text{ wenn } s \longrightarrow 0.$$

Daher ist

$$\begin{aligned} & s^{-2} Q \otimes r \left[\left\{ 2 \left(\frac{dq_s}{dr} \right)^{1/2} - 2 \left(\frac{dq_o}{dr} \right)^{1/2} - s \xi \left(\frac{dq_o}{dr} \right)^{1/2} \right\}^2 \right] \leq \\ & 2 s^{-2} Q \otimes r \left[\left\{ 2 \left(\frac{dq_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_s}{dr} \right)^{1/2} \right\}^2 \right] + \\ & 2 s^{-2} Q \otimes r \left[\left\{ 2 \left(\frac{dp_s}{dr} \right)^{1/2} - 2 \left(\frac{dp_o}{dr} \right)^{1/2} - s \xi \left(\frac{dp_o}{dr} \right)^{1/2} \right\}^2 \right] \longrightarrow 0 \end{aligned}$$

für $s \longrightarrow 0$ als Konsequenz von (3) und der vorausgesetzten

Differenzierbarkeit der Kurve $s \longrightarrow p_s$. Somit ist auch $s \longrightarrow q_s$ $L^2(Q \otimes p_0)$ -differenzierbar. ■

12 Bemerkung

Die Aussage des Lemmas 11 bleibt wahr, wenn die Beziehung (1) ersetzt wird durch

$$(1') \quad s^{-1} d_2(Q \otimes p_s, Q \otimes q_s) \longrightarrow 0 \text{ für } s \longrightarrow 0.$$

Dies wurde im Beweis zu 11 bereits dargetan. ■

Wir wollen nun den Begriff einer differenzierbaren Familie von Markoff-Kernen einführen. Sei dazu $(H, \|\cdot\|)$ ein Banach-Raum und H_0 eine offene Teilmenge von H . Mit

$$p(H_0) := \left\{ p_h : h \in H_0 \right\}$$

bezeichnen wir eine Familie von Markoff-Kernen aus $p(r)$. Wir geben im folgenden zwei Definitionen, die sich als gleichwertig erweisen. Der Grund für die Einführung zweier nahezu gleichwertiger Begriffe liegt vor allem darin, dass die Ableitung bzgl. des einen Differenzierbarkeitsbegriffs genau die relevante Grösse ist bei Anwendungen der Differenzierbarkeit im Zusammenhang mit Untersuchungen zur lokalen asymptotischen Normalität. Die andere Ableitung ist dagegen beweistechnisch leichter zu handhaben.

13 Definition

Sei $h' \in H_0$. Wir nennen $p(H_0)$ in h' $L^2(Q \otimes p_{h'})$ -differenzierbar, falls ein beschränkter linearer Operator

$$Dp_{h'}: H \longrightarrow L^2(Q \otimes p_{h'})$$

existiert mit

$$(1) \quad \|h - h'\|^{-2} Q \otimes p_{h'} \left[\left\{ 2 \left(\frac{dp_h}{dp_{h'}} \right)^{1/2} - 2 - Dp_{h'}(h - h') \right\}^2 \right] \longrightarrow 0$$

$$\text{für } \|h - h'\| \longrightarrow 0 \text{ und}$$

$$(2) \quad \|h - h'\|^{-2} Q \left[p_h(x, N_h(x)) \right] \longrightarrow 0 \text{ für } \|h - h'\| \longrightarrow 0.$$

Dabei ist $\left(\frac{dp_h}{dp_{h'}}, \left\{ N_h(x) : x \in E \right\} \right)$ die Lebesgue-Zerlegung von p_h

bzgl. $p_{h'}$ (vgl. 9.3). Der Operator $Dp_{h'}$ heisst $L^2(Q \otimes p_{h'})$ -Ableitung von $p(H_0)$ in h' . ■

14 Definition

Die Familie von Markoff-Kernen $p(H_0)$ heisst in $h' \in H_0$ $L^2(Q \otimes r)$ -differenzierbar, falls ein beschränkter linearer Operator

$$D'p_{h'} : H \longrightarrow L^2(Q \otimes r)$$

existiert mit

$$\begin{aligned} & \|h - h'\|^{-2} Q \otimes r \left[\left\{ 2 \left(\frac{dp_h}{dr} \right)^{1/2} - 2 \left(\frac{dp_{h'}}{dr} \right)^{1/2} - D'p_{h'}(h - h') \right\}^2 \right] \\ & \longrightarrow 0 \text{ für } \|h - h'\| \longrightarrow 0. \end{aligned}$$

$D'p_{h'}$ nennen wir die $L^2(Q \otimes r)$ -Ableitung von $p(H_0)$ in h' . ■

15 Satz

Folgende zwei Aussagen sind äquivalent.

(1) $p(H_0)$ ist in $h' \in H_0$ $L^2(Q \otimes p_{h'})$ -differenzierbar mit Ableitung $Dp_{h'}$.

(2) $p(H_0)$ ist in $h' \in H_0$ $L^2(Q \otimes r)$ -differenzierbar mit Ableitung $D'p_{h'}$.

Sind (1) oder (2) erfüllt, so gilt für die Ableitungen:

$$D'p_{h'}(h) = Dp_{h'}(h) \left(\frac{dp_{h'}}{dr} \right)^{1/2}.$$

Beweis

Ganz analog zu Satz 9 durch Einsatz einer günstigen Lebesgue-Zerlegung. ■

16 Satz

Wir nehmen an, dass $p(H_0)$ in $h' \in H_0$ $L^2(Q \otimes r)$ -differenzierbar mit Ableitung $D'p_{h'}$ ist. Es sei $s \longrightarrow h_s$ eine in 0 stetige Kurve in H_0 . Dann ist die Kurve

$$s \longrightarrow p_{h'+sh_s} \in p(r)$$

in $p(r)$ $L^2(Q \otimes r)$ -differenzierbar mit Ableitung $s \longrightarrow sD'p_{h'}(h_0)$

und ebenso $L^2(Q \otimes p_h)$ -differenzierbar mit Tangentialvektor $Dp_{h'}(h_0)$.

Beweis

Wir haben

$$s^{-2} Q\otimes r \left[\left\{ 2 \left(\frac{dp_{h'+sh_s}}{dr} \right)^{1/2} - 2 \left(\frac{dp_{h'}}{dr} \right)^{1/2} - s D'p_{h'}(h_o) \right\}^2 \right] \leq$$

$$2 |h_s|^2 |sh_s|^{-2} Q\otimes r \left[\left\{ 2 \left(\frac{dp_{h'+sh_s}}{dr} \right)^{1/2} - 2 \left(\frac{dp_{h'}}{dr} \right)^{1/2} - \right.$$

$$\left. D'p_{h'}(sh_s) \right\}^2 \right] +$$

$$2Q\otimes r \left[\left\{ D'p_{h'}(h_s) - D'p_{h'}(h_o) \right\}^2 \right].$$

Der erste Summand konvergiert infolge der $L^2(Q\otimes r)$ -

Differenzierbarkeit gegen 0. Der zweite verschwindet, weil $D'p_{h'}$ ein beschränkter Operator ist. Der erste Teil der Behauptung ist somit nachgewiesen, der zweite folgt offenbar mit Hilfe von 14. ■

17 Bemerkung

Bei der Ableitung im Sinne der Definition 13 handelt es sich um die Frechet-Ableitung des Operators F

$$F : H_o \longrightarrow L^2(Q\otimes r)$$

$$F(h) := 2 \left(\frac{dp_h}{dr} \right)^{1/2}.$$

an der Stelle h' . Die Ableitungen der Kurven $s \longrightarrow p_{h'+sh}$ entsprechen der Richtungsableitung des Operators F in h' in Richtung von h . Um mit Hilfe der Existenz differenzierbarer Kurven auf die Differenzierbarkeit einer Familie schliessen zu können, kann man sich also bekannter Sätze der Funktionalanalysis bedienen. ■

Für den Fall einer offenen und konvexen Teilmenge $H_o \subseteq \mathbb{R}^k$ geben wir folgendes Differenzierbarkeitskriterium an, welches in Anlehnung an die entsprechende Aussage für Verteilungen in Witting, 1.194 formuliert wird. Wir nennen zwei Markoff-Kerne p und q äquivalent, wenn für jedes $x \in E$ die Masse $p(x, \cdot)$ und $q(x, \cdot)$ äquivalent sind.

18 Satz

Ist $H_0 \subseteq \mathbb{R}^k$ offen und konvex, $h' \in H_0$ und $p(H_0)$ eine Familie paarweise äquivalenter Markoff-Kerne, so ist $p(H_0)$ in h' $L^2(Q_{\text{sep}_{h'}})$ -differenzierbar mit Ableitung

$$(1) D_{p_{h'}}(h) = h^t \left[D_h \log \left(\frac{dp_h}{dr} \right) \Big|_{h=h'} \right],$$

falls gilt:

$$(2) D_h \left(\frac{dp_h}{dr} \right)(x,y) \text{ existiert und ist stetig für } Q_{\text{sep}_{h'}}\text{-f.a. } (x,y).$$

$$(3) \frac{\partial}{\partial h_i} \log \frac{dp_h}{dr} \in L^2(Q_{\text{sep}_h}) \text{ für alle } i = 1, \dots, k.$$

$$(4) Q_{\text{sep}_h} \left[\left\{ \frac{\partial}{\partial h_i} \log \left(\frac{dp_h}{dr} \right) \right\}^2 \right] \longrightarrow Q_{\text{sep}_{h'}} \left[\left\{ \frac{\partial}{\partial h_i} \log \left(\frac{dp_h}{dr} \right) \Big|_{h=h'} \right\}^2 \right]$$

für $|h - h'| \longrightarrow 0$ und alle $i = 1, \dots, k$.

Beweis

Infolge der paarweisen Äquivalenz der beteiligten Markoff-Kerne ist die Bedingung (2) aus 13 erfüllt. Da nun $p(H_0)$ dann und nur dann in h' $L^2(Q_{\text{sep}_{h'}})$ -differenzierbar ist mit Ableitung wie unter (1) angegeben, wenn die Familie $\{Q_{\text{sep}_h} : h \in H_0\}$ von Verteilungen auf $(E^2, \mathcal{B}(E^2))$ differenzierbar im Sinne der Definition Witting, 1.187 ist, so ergibt sich die Behauptung aus Satz 1.194, Witting.

■

19 Bemerkung

Wir wollen illustrieren, wie 'nahe' die Voraussetzungen des Satzes 18 an seiner Behauptung liegen. Dazu betrachten wir einen endlichen Zustandsraum $E := \{1, \dots, N\}$, eine offene und konvexe Teilmenge H_0 von $\mathbb{R}^k$ und eine Familie $p(H_0)$ von Markoff-Kernen auf $(E, \mathcal{B}(E))$. Mit $r(i, \cdot)$ bezeichnen wir für jedes $i \in E$ das Zählmaß auf $\mathcal{B}(E)$. Dann ist

$$\frac{dp_h}{dr}(i, j) = p_h(i, \cdot)(j) =: p_h(i, j).$$

Wir nehmen an, dass $p_h(i, j) > 0$ für alle $(i, j) \in E^2$. Q sei eine Verteilung auf $\mathcal{B}(E)$ mit $Q(i) > 0$ für alle $i \in E$. Für alle h aus

einer Umgebung von $h' \in H_0$ gelte:

$p(H_0)$ ist in $L^2(Q \otimes p_h)$ -differenzierbar mit Ableitung Dp_h , und die Abbildungen

$$h \longrightarrow Dp_h(e_l) : H_0 \longrightarrow L^2(Q \otimes p_h)$$

sind stetig in h' für jeden Vektor e_l aus der Standard-Normalbasis des $\mathbb{R}^k$. Dann sind (1) bis (4) aus Satz 18 erfüllt.

Aus $Q \otimes p_h(i, j) > 0$ für alle i und $j \in E$ und der $L^2(Q \otimes p_h)$ -Differenzierbarkeit von $p(H_0)$ in einer Umgebung von h' folgt unter Berücksichtigung des Satzes 15, dass punktweise für jedes $(i, j) \in E^2$ gilt:

$$(2) \quad |g - h|^{-1} \left\{ 2p_g^{1/2} - 2p_h^{1/2} - Dp_h(g - h)p^{1/2} \right\} \longrightarrow 0$$

für $g \longrightarrow h$. Also ist für feste (i, j) die Abbildung

$$(3) \quad h \longrightarrow 2p_h^{1/2}(i, j) : \mathbb{R}^k \longrightarrow \mathbb{R}$$

total differenzierbar im klassischen analytischen Sinne mit Jacobi-Matrix

$$(4) \quad \left[\frac{\partial 2p_h^{1/2}(i, j)}{\partial h_l} \right] = Dp_h(e_l)(i, j).$$

Hieraus sowie aus (2) ergeben sich leicht die Beziehungen (1) bis (4) aus 18. ■

20 Bemerkung

Die hier eingeführten Differenzierbarkeitsbegriffe sind mit den für Verteilungen bereits bekannten konsistent. Genauer gilt folgendes. Bezeichne R ein σ -finites Mass auf $(E, \mathcal{B}(E))$ und $P(R)$ die Menge aller Verteilungen P auf $(E, \mathcal{B}(E))$, die eine Dichte bzgl. R haben. Ist $H_0 \subseteq \mathbb{R}^k$ offen, so sei $P(H_0) \subseteq P(R)$ eine Familie von Verteilungen auf $(E, \mathcal{B}(E))$. Jedes dieser Masse können wir auch als Kern auf $(E, \mathcal{B}(E))$ auffassen (der von seinem ersten Argument nicht abhängt). Wie oben sei Q eine beliebige Verteilung auf $(E, \mathcal{B}(E))$. Dann gilt

(1) Eine Kurve $s \longrightarrow P_s$ ist $L^2(Q \otimes R)$ -differenzierbar mit Ableitung $s \longrightarrow s_h$ dann und nur dann, wenn sie differenzierbar ist im Sinne der Definition aus Strasser (1), 75. mit Ableitung

$$s \longrightarrow \frac{s}{2} \hat{h},$$

wobei $\hat{h} := Q(h(\cdot, y))$.

(2) Genau dann ist die Familie $P(H_0)$ in $h' \in H_0$ $L^2(Q \otimes P_{h'})$ -differenzierbar mit Ableitung $DP_{h'}$, wenn sie differenzierbar im Sinne der Definition Witting, 1.187 ist mit Ableitung

$$L_{h',h}(y) = Q(DP_{h'}(h)(\cdot, y)).$$

(3) $P(H_0)$ ist in $h' \in H_0$ genau dann $L^2(Q \otimes R)$ -differenzierbar mit Ableitung $D'P_{h'}$, wenn sie differenzierbar im Sinne der Definition aus Strasser (1), 76. ist mit Ableitung

$$h \longrightarrow Q(D'P_{h'}(h)(\cdot, y)).$$

Beweis

Wir führen nur den Beweis der ersten Behauptung exemplarisch vor.

Sei die Kurve $L^2(Q \otimes R)$ -differenzierbar mit Ableitung $s \longrightarrow sh$.

Dann ist mit Hilfe der Jensen-Ungleichung und des Satzes von

Fubini

$$\begin{aligned} & |s|^{-2} R \left[\left\{ \left(\frac{dP_s}{dR} \right)^{1/2} (y) - \left(\frac{dP_0}{dR} \right)^{1/2} (y) - \frac{s}{2} Q(h(\cdot, y)) \right\}^2 \right] = \\ & \frac{1}{4} |s|^{-2} R \left[\left\{ Q \left[2 \left(\frac{dP_s}{dR} \right)^{1/2} - 2 \left(\frac{dP_0}{dR} \right)^{1/2} - sh \right] \right\}^2 \right] \leq \\ & \frac{1}{4} |s|^{-2} Q \otimes R \left[\left\{ 2 \left(\frac{dP_s}{dR} \right)^{1/2} - 2 \left(\frac{dP_0}{dR} \right)^{1/2} - sh \right\}^2 \right] \longrightarrow 0. \end{aligned}$$

Daraus folgt die Differenzierbarkeit der Kurve im Sinne der Definition aus Strasser (1). Umgekehrt nehmen wir an, die Kurve sei im Sinne dieser Definition differenzierbar und die Ableitung sei $s \longrightarrow sh$ für ein $h \in L^2(R)$. Dann ist $h \in L^2(Q \otimes R)$ und die $L^2(Q \otimes R)$ -Differenzierbarkeit folgt sogleich, weil die zu bildenden Integrale vom Mass Q gar nicht abhängen. ■

21 Beispiel

Betrachte einen endlichen Zustandsraum $E = \{1, \dots, N\}$ versehen mit seiner Potenzmenge als σ -Algebra $\mathcal{B}(E)$. Jeder Markoff-Kern auf $(E, \mathcal{B}(E))$ besitzt eine Dichte bzgl. des Zählkerns

$$r(x, B) := \sum_{k=1}^N \varepsilon_k(B) \quad \text{für alle } x \in E.$$

Eine Kurve von Markoff-Kernen p_s auf $(E, \mathcal{B}(E))$ ist $L^2(Q \otimes r)$ -differenzierbar dann und nur dann, wenn die Abbildungen

$$s \longrightarrow p_s(x, \langle y \rangle)$$

für beliebige, feste (x, y) aus dem Träger des Masses $Q \otimes p_0$ differenzierbar im Sinne der Analysis sind. ■

In manchen Situationen ist der Parameterraum einer Familie von Markoff-Kernen in natürlicher Weise durch eine differenzierbare Mannigfaltigkeit gegeben. Wir wollen nun auch für diesen Fall einen Differenzierbarkeitsbegriff einführen. Er stellt eine Verallgemeinerung der bisher geprägten Begriffe dar. Die hier zugrundeliegende Definition einer differenzierbaren Mannigfaltigkeit weicht von der in der Literatur üblicherweise anzutreffenden ab (vgl. 10, insbesondere 10.6). Er ist allerdings eine Spezialisierung dessen. Der Grund für die Wahl unserer Begriffsbildung ist der, dass wir die geometrischen Eigenschaften des Grundraumes der Mannigfaltigkeit stärker berücksichtigen wollen.

Für den Rest dieses Abschnitts sei B ein Banach-Raum, H ein separabler Hilbert-Raum und M eine differenzierbare Hilbert-Mannigfaltigkeit in B , also $M \subseteq B$. Karten von M notieren wir mit (U, Ψ, V) . Dabei ist U eine relativ offene Menge in M , V eine offene Menge in H und Ψ ein Diffeomorphismus von U auf V . Sei

$$p(M) := \left\{ p_m : m \in M \right\} \subseteq p(r)$$

eine Familie von Markoff-Kernen auf dem polnischen Raum $(E, \mathcal{B}(E))$.

22 Definition

Ist $m' \in M$, so heisst $p(M)$ in m' $L^2(Q \otimes r)$ -differenzierbar, falls ein beschränkter linearer Operator

$$D'p_{m'} : T_{\xi_{m'}} M \longrightarrow L^2(Q \otimes r)$$

existiert, so dass für eine beliebige Karte (U, Ψ, V) von M mit $m' \in U$ - und damit für jede solche Karte - gilt

$$(1) \quad |m - m'|^{-2} \otimes r \left[\left\{ 2 \left(\frac{dp_m}{dr} \right)^{1/2} - 2 \left(\frac{dp_{m'}}{dr} \right)^{1/2} - D'p_{m'} \left\{ \left(D\Psi_{m'} \right)^{-1} \left(\Psi(m) - \Psi(m') \right) \right\} \right\}^2 \right] \longrightarrow 0$$

für $|m - m'| \longrightarrow 0$. ■

23 Bemerkung

Die Unabhängigkeit der Definition von der Kartenwahl sieht man so ein: Ist (U', Ψ', V') eine weitere Karte mit $m' \in U'$ und bezeichnen wir die gemäss der Definition der Differenzierbarkeit der Abbildungen Ψ und Ψ' existierenden Fortsetzungen von Ψ und Ψ' auf offene Teilmengen von $\mathbb{B}$ mit dem gleichen Symbol, so ist

$$\begin{aligned} & \left| D'p_{m'} \left\{ \left(D\Psi_{m'} \right)^{-1} \left(\Psi(m) - \Psi(m') \right) \right\} - \right. \\ & \left. D'p_{m'} \left\{ \left(D\Psi'_{m'} \right)^{-1} \left(\Psi'(m) - \Psi'(m') \right) \right\} \right| \leq \\ & |D'p_{m'}| \left| \left(D\Psi_{m'} \right)^{-1} \left(D\Psi_{m'}(m - m') + o(|m - m'|) \right) - \right. \\ & \left. \left(D\Psi'_{m'} \right)^{-1} \left(D\Psi'_{m'}(m - m') + o(|m - m'|) \right) \right| = \\ & o(|m - m'|). \end{aligned}$$

Hieraus folgt die Behauptung. ■

24 Bemerkung

Genau dann ist $p(M)$ in $m' \in L^2(Q \otimes r)$ -differenzierbar, wenn

$$q(V) := \left\{ q_h := p_{\Psi^{-1}(h)} : h \in V \right\}$$

in $h' := \Psi(m') \in L^2(Q \otimes r)$ -differenzierbar ist im Sinne der Definition

14 für jede beliebige Karte (U, Ψ, V) von M mit $m' \in U$. Für die Ableitungen gilt

$$D'p_{m'} = D'q_{h'} \circ D\Psi_{m'}. \quad \blacksquare$$

25 Satz und Definition

$p(M)$ ist in $m' \in M$ $L^2(Q \otimes r)$ -differenzierbar mit Ableitung $D'p_{m'}$ dann und nur dann, wenn sie $L^2(Q \otimes p_{m'})$ -differenzierbar ist mit

Ableitung $Dp_{m'}$, d.h. falls ein beschränkter linearer Operator

$$Dp_{m'} : Tg_{m'}M \longrightarrow L^2(Q \otimes p_{m'})$$

existiert, für den folgende zwei Bedingungen mit einer - und damit jeder Karte (U, Ψ, V) von M mit $m' \in U$ - erfüllt sind:

$$(1) |m - m'|^{-2} \quad Q \otimes p_{m'} \left[\left\{ 2 \left(\frac{dp_m}{dp_{m'}} \right)^{1/2} - 2 - Dp_{m'} \left\{ \left(D\Psi_{m'} \right)^{-1} \left(\Psi(m) - \Psi(m') \right) \right\} \right\}^2 \right] \longrightarrow 0$$

für $|m - m'| \longrightarrow 0$ und

$$(2) |m - m'|^{-2} Q(p_m(x), N_m(x)) \longrightarrow 0$$

für $|m - m'| \longrightarrow 0$ und die singulären Anteile $N_{m'}$ aus der Lebesgue-Zerlegung von p_m bzgl. $p_{m'}$. In diesem Fall besteht folgende Gleichheit:

$$Dp_{m'} \left(\frac{dp_{m'}}{dr} \right)^{1/2} = D'p_{m'}$$

Beweis

Diese Aussage ist eine Konsequenz der Bemerkung 24 und des Satzes 15. ■

26 Satz

Sei $s \longrightarrow m_s \in M$ eine Kurve mit

$$s^{-1}(m_s - m') \longrightarrow t \in Tg_{m'}M$$

für $s \longrightarrow 0$. Wenn $p(M)$ in $m' \in M$ $L^2(Q \otimes p_{m'})$ -differenzierbar ist mit Ableitung $Dp_{m'}$, so ist die Kurve $s \longrightarrow p_{m_s} \in p(r)$ $L^2(Q \otimes r)$ -differenzierbar mit Tangentialvektor $Dp_{m'}(t)$.

Beweis

(U, Ψ, V) sei eine Karte von M und $m' \in U$. Bezeichne $h_s := \Psi(m_s)$.

Dann gilt

$$\begin{aligned}
& |s|^{-2} \text{Qor} \left[\left\{ 2 \left(\frac{dp_{m'}}{dr} \right)^{1/2} - 2 \left(\frac{dp_m}{dr} \right)^{1/2} \right. \right. \\
& \quad \left. \left. \leq Dp_{m'}(t) \left(\frac{dp_{m'}}{dr} \right)^{1/2} \right\}^2 \right] = \\
& |s|^{-2} \text{Qor} \left[\left\{ 2 \left(\frac{dq_{h_s}}{dr} \right)^{1/2} - 2 \left(\frac{dq_{h'}}{dr} \right)^{1/2} \right. \right. \\
& \quad \left. \left. \leq Dq_{h_s} \circ D\Psi_{m'}(t) \left(\frac{dq_{h'}}{dr} \right)^{1/2} \right\}^2 \right].
\end{aligned}$$

Dabei ist $q_{h_s} := p_{\Psi^{-1}(h_s)}$. Nun haben wir

$$\begin{aligned}
h_s &= h' + s \cdot s^{-1}(h_s - h') \\
\text{und} \\
s^{-1}(h_s - h') &= s^{-1} \left[D\Psi_{m'}(m_s - m') + o(|m_s - m'|) \right] \longrightarrow \\
D\Psi_{m'}(t) \quad \text{für } |m_s - m'| &\longrightarrow 0.
\end{aligned}$$

Hieraus ergibt sich mit Hilfe des Satzes 16 die Behauptung. ■

27 Satz

Betrachte eine differenzierbare Untermannigfaltigkeit $M^1 \subseteq M$ von M und die Familie

$$p(M^1) := \left\{ p_m : m \in M^1 \right\}.$$

Sei $m' \in M^1$ und $p(M)$ in $m' \in L^2(\text{Qor})$ -differenzierbar. Dann ist auch $p(M^1)$ in $m' \in L^2(\text{Qor})$ -differenzierbar und die Ableitung ist

$$D'p_{m'} \Big|_{T\mathfrak{L}_{m'} M^1}.$$

Beweis

Wir geben uns zwei Karten (U, Ψ, V) und (U^1, Ψ^1, V^1) von M bzw. M^1 vor mit $m' \in U \cap U^1$. Dann gilt für $m \in M^1$

$$\begin{aligned}
& |m - m'|^{-2} \text{Qor} \left[\left\{ 2 \left(\frac{dp_m}{dr} \right)^{1/2} - 2 \left(\frac{dp_{m'}}{dr} \right)^{1/2} \right. \right. \\
& \quad \left. \left. D'p_{m'} \left\{ \left(D\Psi_{m'} \right)^{-1} \left(\Psi^1(m) - \Psi^1(m') \right) \right\} \right\}^2 \right] \leq
\end{aligned}$$

$$\begin{aligned}
& 2|m - m'|^{-2} \operatorname{Qor} \left[\left\{ 2 \left(\frac{dp_m}{dr} \right)^{1/2} - 2 \left(\frac{dp_{m'}}{dr} \right)^{1/2} - \right. \right. \\
& \quad \left. \left. D'p_{m'} \left\{ \left(D\Psi_{m'} \right)^{-1} \left(\Psi(m) - \Psi(m') \right) \right\} \right\}^2 + \right. \\
& \quad \left. 2|m - m'|^{-2} \operatorname{Qor} \left[\left\{ D'p_m \left\{ \left(D\Psi_m^1 \right)^{-1} \left(\Psi^1(m) - \Psi^1(m') \right) \right\} \right\} - \right. \right. \\
& \quad \left. \left. D'p_{m'} \left\{ \left(D\Psi_{m'} \right)^{-1} \left(\Psi(m) - \Psi(m') \right) \right\} \right\}^2 \right].
\end{aligned}$$

Der erste Summand konvergiert hier infolge der Differenzierbarkeit der Familie $p(M)$ in m' gegen 0. Der zweite verschwindet asymptotisch, weil $D'p_{m'}$ ein stetiger Operator ist und

$$\begin{aligned}
& |m - m'|^{-1} \left\{ D'p_{m'} \left\{ \left(D\Psi_{m'}^1 \right)^{-1} \left(\Psi^1(m) - \Psi^1(m') \right) \right\} \right\} - \\
& \quad D'p_{m'} \left\{ \left(D\Psi_{m'} \right)^{-1} \left(\Psi(m) - \Psi(m') \right) \right\} \Bigg\}
\end{aligned}$$

gegen 0 strebt. ■

7. Stochastische Entwicklung von Dichtequotienten

Wir gehen aus von einem polnischen Raum $(E, \mathcal{B}(E))$. Für jedes $n \in \mathbb{N}$ seien auf ihm Markoff-Kerne

$$q_{n,1}, \dots, q_{n,n-1} \quad \text{und} \quad p_{n,1}, \dots, p_{n,n-1}$$

sowie Kerne

$$r_{n,1}, \dots, r_{n,n-1}$$

gegeben. Wir nehmen an, dass Dichten

$$\frac{dq_{n,i}}{dr_{n,i}} \quad \text{und} \quad \frac{dp_{n,i}}{dr_{n,i}}, \quad 1 \leq i \leq n-1, \quad n \geq 2,$$

existieren. Dann sind auch die Lebesgue-Zerlegungen

$$\left[\frac{dq_{n,i}}{dp_{n,i}}, \left\{ N_{n,i}(x) : x \in E \right\} \right]$$

definiert (vgl. 9.5). Für jedes $n \in \mathbb{N}$ definiere einen Markoff-Kern von $(E, \mathcal{B}(E))$ nach $(E^{n-1}, \mathcal{B}(E^{n-1}))$ durch

$$q_n(x, B_2 x \dots x B_n) :=$$

$$\int_{B_2} q_{n,1}(x, dx_2) \int_{B_3} q_{n,2}(x_2, dx_3) \dots \int_{B_n} q_{n,n-1}(x_{n-1}, dx_n).$$

Entsprechend seien p_n und r_n definiert. Dann existieren abermals Dichten

$$\frac{dq_n}{dr_n} \quad \text{und} \quad \frac{dp_n}{dr_n}$$

und auch die Lebesgue-Zerlegung

$$\left[\frac{dq_n}{dp_n}, \left\{ N_n(x) : x \in E \right\} \right]$$

(vgl. 9.13). $Q_{n,1}$ und $P_{n,1}$ seien zwei Verteilungen auf $(E, \mathcal{B}(E))$.

Die Lebesgue-Zerlegung von $Q_{n,1}$ bzgl. $P_{n,1}$ bezeichnen wir mit

$$\left[\frac{dQ_{n,1}}{dP_{n,1}}, M_{n,1} \right].$$

Für $n \in \mathbb{N}$ definieren wir die Verteilung Q_n auf $(E^n, \mathcal{B}(E^n))$ durch

$$Q_n \left[B_1 x \dots x B_n \right] := Q_{n,1} \left[1_{B_1}(x_1) q_n(x_1, B_2 x \dots x B_n) \right],$$

$B_1, \dots, B_n \in \mathcal{B}(E)$. Analog sei die Verteilung P_n definiert. Für $i \in \mathbb{N}$, $i < n$ setze

$$Q_{n,i,i+1} \left\{ \begin{matrix} B_1 \times B_2 \end{matrix} \right\} := Q_n \left\{ \begin{matrix} C_1 \times \dots \times C_n \end{matrix} \right\}$$

mit

$$C_j := \begin{cases} B_1 & \text{für } j = 1 \\ B_2 & \text{für } j = i + 1 \\ E & \text{sonst} \end{cases}$$

und

$$Q_{n,i}(B) := Q_{n,i,i+1}(B \times E).$$

Entsprechende Grössen sollen für $P_{n,i}$ und p_n definiert sein.

Die Lebesgue-Zerlegung von Q_n bzgl. P_n (vgl. 9.14) notieren wir schliesslich mit

$$\left(\frac{dQ_n}{dP_n}, M_n \right).$$

In diesem Abschnitt werden wir eine stochastische Entwicklung der

Dichtequotienten $\frac{dQ_n}{dP_n}$ angeben. Sie erweist sich bei unseren

Untersuchungen zur lokalen asymptotischen Normalität als ein wesentliches Werkzeug. Eine solche stochastische Entwicklung ist

stets möglich, wenn gewisse Funktionen der Dichtequotienten $\frac{dQ_{n,i}}{dP_{n,i}}$

Eigenschaften aufweisen, die in der folgenden Definition benannt werden.

1 Definition

Eine Familie $\left\{ \xi_{n,i} : 1 \leq i \leq n-1, n > 1 \right\}$ von $L^2(P_{n,i,i+1})$ -Funktionen heisst Lindeberg-Feld für

$$\left\{ \left(E^n, \mathcal{B}(E^n), P_n \right) : n \in \mathbb{N} \right\}$$

oder, sofern der Zusammenhang klar ist, kurz Lindeberg-Feld, wenn

$$(1) \quad \limsup_{n \rightarrow \infty} \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\xi_{n,i}^2 \right] < \infty$$

und

$$(2) \quad \lim_{n \rightarrow \infty} \sum_{l=1}^{n-1} P_{n,l,l+1} \left[\xi_{n,l}^2 \mathbb{1}_{\left\{ |\xi_{n,l}| > \varepsilon \right\}} \right] = 0$$

für alle $\varepsilon > 0$.

Es gilt nun folgender Entwicklungssatz.

2 Satz

Ist für $n > 1$ und $1 \leq i < n$

$$(1) \quad \xi_{n,l} := 2 \left[\frac{dq_{n,l}}{dp_{n,l}} \right]^{1/2} - 2$$

und

(2) $\left\{ \xi_{n,l} : 1 \leq l \leq n-1, n > 1 \right\}$ ein Lindeberg-Feld für

$\left\{ \left(E^n, \mathcal{E}(E^n), P_n \right) : n \in \mathbb{N} \right\}$ so gilt:

$$(3) \quad \frac{dQ_n}{dP_n}(\underline{x}) = \frac{dQ_{n,1}}{dP_{n,1}}(x_1) \exp \left\{ \sum_{l=1}^{n-1} \xi_{n,l} \langle x_l, x_{l+1} \rangle - \sum_{l=1}^{n-1} P_{n,l} \left[x_l, \xi_{n,l} \langle x_l, \cdot \rangle \right] - \frac{1}{2} \sum_{l=1}^{n-1} P_{n,l} \left[x_l, \xi_{n,l}^2 \langle x_l, \cdot \rangle \right] - \sum_{l=1}^{n-1} q_{n,l} \left[x_l, N_{n,l} \langle x_l \rangle \right] + o_{P_n}(n^0) \right\}.$$

Beweis

Wir haben

$$(4) \quad \frac{dQ_n}{dP_n}(\underline{x}) = \frac{dQ_{n,1}}{dP_{n,1}}(x_1) \prod_{i=1}^{n-1} \frac{dq_{n,i}}{dp_{n,i}}(x_i, x_{i+1})$$

für $\underline{x} \notin N_n$ mit

$$(5) \quad N_n = \left\{ \underline{x} \in E^n : \frac{dP_{n,1}}{d(P_{n,1} + Q_{n,1})}(x_1) \prod_{i=1}^{n-1} \frac{dp_{n,i}}{dr_{n,i}}(x_i, x_{i+1}) = 0 \right\}.$$

Weil $P_n(N_n) = 0$, gilt (4) P_n -f.s. und somit

$$(6) \quad \frac{dQ_n}{dp_n}(\underline{x}) = \frac{dQ_{n,1}}{dp_{n,1}}(\underline{x}_1) \prod_{i=1}^{n-1} \frac{dq_{n,i}}{dp_{n,i}}(x_i, x_{i+1}) \exp \left[\phi_{p_n}(n^0)(\underline{x}) \right]$$

Nun ist

$$\frac{dq_{n,i}}{dp_{n,i}} = \frac{1}{2} \left(\xi_{n,i} + 1 \right)^2$$

und deshalb

$$P_n \left\{ \frac{dq_n}{dp_n}(\underline{x}) = 0 \right\} =$$

$$P_n \left\{ \frac{dq_{n,i}}{dp_{n,i}}(x_i, x_{i+1}) = 0 \text{ für ein } i \in \{1, \dots, n-1\} \right\} =$$

$$P_n \bigcup_{i=1}^{n-1} \left\{ \xi_{n,i}(x_i, x_{i+1}) = -2 \right\} \leq$$

$$\sum_{i=1}^{n-1} P_{n,i,i+1} \left\{ |\xi_{n,i}| \geq 2 \right\} \leq$$

$$\sum_{i=1}^{n-1} P_{n,i,i+1} \left[\xi_{n,i}^2 \mathbb{1}_{\left\{ |\xi_{n,i}| > 1 \right\}} \right]$$

Mit Hilfe von 1(2) folgt nun

$$P_n \left\{ \frac{dq_n}{dp_n}(\underline{x}) = 0 \right\} \longrightarrow 0.$$

Darum genügt es, solche $\underline{x} \in E^n$ zu betrachten, für die

$$(7) \quad \frac{dq_n}{dp_n}(\underline{x}) > 0$$

oder gleichwertig

$$(8) \quad \xi_{n,i}(x_i, x_{i+1}) > -2 \text{ für alle } i = 1, \dots, n-1$$

erfüllt ist. Gemäss 11.4 ist für $s > -2$

$$2 \log \left[\frac{s}{2} + 1 \right] = s - \frac{s^2}{4} + \frac{s^2}{4} (1 - r(s))$$

mit einer in 0 gleichmässig stetigen Funktion r . Für jedes $\underline{x} \in E^n$, welches (8) genügt, gilt daher

$$(9) \quad \log \frac{dq_n}{dp_n} = 2 \log \left[\prod_{i=1}^{n-1} \left[\frac{1}{2} \xi_{n,i} \langle x_i, x_{i+1} \rangle + 1 \right] \right] = \\ \sum_{i=1}^{n-1} \left[\xi_{n,i} \langle x_i, x_{i+1} \rangle - \frac{1}{4} \xi_{n,i}^2 \langle x_i, x_{i+1} \rangle + \right. \\ \left. \frac{1}{4} \xi_{n,i}^2 \langle x_i, x_{i+1} \rangle \left\{ 1 - r \left(\xi_{n,i} \langle x_i, x_{i+1} \rangle \right) \right\} \right].$$

Ist $\varepsilon > 0$ eine reelle Zahl, so ist

$$P_n \left\{ \max \left[\left| \xi_{n,i} \langle x_i, x_{i+1} \rangle \right| : 1 \leq i < n \right] > \varepsilon \right\} \leq \\ \sum_{i=1}^{n-1} P_n \left\{ \left| \xi_{n,i} \langle x_i, x_{i+1} \rangle \right| > \varepsilon \right\} \leq \\ \varepsilon^{-2} \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\xi_{n,i}^2 \mathbb{1}_{\left\{ \left| \xi_{n,i} \langle x_i, x_{i+1} \rangle \right| > 1 \right\}} \right].$$

Gemäss 1(2) ist folglich

$$(10) \quad P_n \left\{ \max \left[\left| \xi_{n,i} \langle x_i, x_{i+1} \rangle \right| : 1 \leq i < n \right] > \varepsilon \right\} \longrightarrow 0$$

und darum

$$(11) \quad \max \left[1 - r(\xi_{n,i}) : 1 \leq i < n \right] = o_{P_n}(n^0).$$

Ausserdem ist die Folge $\left[\sum_{i=1}^{n-1} \xi_{n,i}^2 \langle x_i, x_{i+1} \rangle : n \in \mathbb{N} \right]$ wegen 1(1)

gemäss 11.5 (3) P_n -stochastisch beschränkt. Daher erhalten wir mit Hilfe von 11.5 (2) und (11)

$$(12) \quad \sum_{i=1}^{n-1} \frac{1}{4} \xi_{n,i}^2 \langle x_i, x_{i+1} \rangle \left\{ 1 - r \left(\xi_{n,i} \langle x_i, x_{i+1} \rangle \right) \right\} = o_{P_n}(n^0).$$

Der rechte Ausdruck auf der rechten Seite von (9) verschwindet also in P_n -Wahrscheinlichkeit. Es ist

$$(13) \quad \sum_{i=1}^{n-1} \left\{ \xi_{n,i}^2 \langle x_i, x_{i+1} \rangle - p_{n,i} \left[x_i, \xi_{n,i}^2 \langle x_i, \cdot \rangle \right] \right\} = o_{P_n}(n^0).$$

Um dies nachzuweisen, definiere für ein beliebig vorgegebenes $\delta > 0$

$$\bar{\xi}_{n,i} := \xi_{n,i} \mathbb{1}_{\left\{ \left| \xi_{n,i} \right| \leq \delta \right\}}.$$

Dann ist

$$\begin{aligned}
& P_n \left[\left| \sum_{i=1}^{n-1} \left\{ \xi_{n,i}^2(x_i, x_{i+1}) - p_{n,i} \left[x_i, \xi_{n,i}^2(x_i, \cdot) \right] \right\} \right| > \varepsilon \right] \leq \\
& P_n \left[\left| \sum_{i=1}^{n-1} \left\{ \xi_{n,i}^2(x_i, x_{i+1}) - \bar{\xi}_{n,i}^2(x_i, x_{i+1}) \right\} \right| > \frac{\varepsilon}{3} \right] + \\
& P_n \left[\left| \sum_{i=1}^{n-1} \left\{ \bar{\xi}_{n,i}^2(x_i, x_{i+1}) - p_{n,i} \left[x_i, \bar{\xi}_{n,i}^2(x_i, \cdot) \right] \right\} \right| > \frac{\varepsilon}{3} \right] + \\
& P_n \left[\left| \sum_{i=1}^{n-1} p_{n,i} \left[x_i, \bar{\xi}_{n,i}^2(x_i, \cdot) - \xi_{n,i}^2(x_i, \cdot) \right] \right| > \frac{\varepsilon}{3} \right].
\end{aligned}$$

Für den ersten Summanden gilt

$$\begin{aligned}
& P_n \left[\left| \sum_{i=1}^{n-1} \left\{ \xi_{n,i}^2(x_i, x_{i+1}) - \bar{\xi}_{n,i}^2(x_i, x_{i+1}) \right\} \right| > \frac{\varepsilon}{3} \right] \leq \\
& \frac{3}{\varepsilon} \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\xi_{n,i}^2 \mathbf{1}_{\left\{ |\xi_{n,i}| > \delta \right\}} \right] \longrightarrow 0
\end{aligned}$$

für $n \longrightarrow \infty$, für alle $\varepsilon > 0$ und alle $\delta > 0$ gemäss 1 (2). Der dritte Summand erfüllt ebenfalls

$$\begin{aligned}
& P_n \left[\left| \sum_{i=1}^{n-1} p_{n,i} \left[x_i, \bar{\xi}_{n,i}^2(x_i, \cdot) - \xi_{n,i}^2(x_i, \cdot) \right] \right| > \frac{\varepsilon}{3} \right] \leq \\
& \frac{3}{\varepsilon} \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\xi_{n,i}^2 \mathbf{1}_{\left\{ |\xi_{n,i}| > \delta \right\}} \right] \longrightarrow 0.
\end{aligned}$$

Schliesslich schätzen wir den zweiten Summanden ab. Es ist

$$\begin{aligned}
& P_n \left[\left| \sum_{i=1}^{n-1} \left\{ \bar{\xi}_{n,i}^2(x_i, x_{i+1}) - p_{n,i} \left[x_i, \bar{\xi}_{n,i}^2(x_i, \cdot) \right] \right\} \right| > \frac{\varepsilon}{3} \right] \leq \\
& \frac{9}{\varepsilon^2} P_n \left[\left(\sum_{i=1}^{n-1} \left\{ \bar{\xi}_{n,i}^2(x_i, x_{i+1}) - p_{n,i} \left[x_i, \bar{\xi}_{n,i}^2(x_i, \cdot) \right] \right\} \right)^2 \right] = \\
& \frac{9}{\varepsilon^2} \sum_{i=1}^{n-1} P_n \left[\left(\bar{\xi}_{n,i}^2(x_i, x_{i+1}) - p_{n,i} \left[x_i, \bar{\xi}_{n,i}^2(x_i, \cdot) \right] \right)^2 \right].
\end{aligned}$$

Diese Gleichheit gilt laut 8.3. Der letzte Ausdruck ist gleich

$$\frac{9}{\varepsilon^2} \sum_{i=1}^{n-1} P_{n,i} \left[p_{n,i} \left[x_i, \left(\bar{\xi}_{n,i}^2(x_i, \cdot) - p_{n,i} \left[x_i, \bar{\xi}_{n,i}^2(x_i, \cdot) \right] \right)^2 \right] \right].$$

Gemäss 8.4 stimmt dieser Ausdruck überein mit

$$\frac{9}{\varepsilon^2} \sum_{i=1}^{n-1} \left\{ P_{n,i} \left[P_{n,i} \left[x_i, \xi_{n,i}^{-4}(x_{i,\cdot}) \right] - \left[P_{n,i} \left[P_{n,i} \left[x_i, \xi_{n,i}^{-2}(x_{i,\cdot}) \right] \right]^2 \right] \right\} \leq \frac{9}{\varepsilon^2} \delta^2 \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\xi_{n,i}^2 \right].$$

Wegen 1 (1) ist der Summenfaktor im Ausdruck oben beschränkt. Deshalb strebt der gesamte Ausdruck mit $\delta \rightarrow 0$ ebenfalls zur 0, und zwar für beliebige $\varepsilon > 0$. Damit ist (13) bewiesen, und hieraus erhalten wir zusammen mit (12) und (9)

$$(14) \quad \frac{dQ_n}{dP_n}(x) = \exp \left\{ \sum_{i=1}^{n-1} \xi_{n,i}(x_i, x_{i+1}) - \frac{1}{4} \sum_{i=1}^{n-1} P_{n,i} \left[x_i, \xi_{n,i}^2(x_{i,\cdot}) \right] + o_{P_n}(n^0) \right\}.$$

Nach bekannten Aussagen über den Hellinger-Abstand zweier Wahrscheinlichkeitsmasse gilt (vgl. 11.2 (5) und 11.2 (6))

$$(15) \quad p_{n,i} \left[x_i, \xi_{n,i}(x_{i,\cdot}) \right] = -2d_2^2 \left[p_{n,i}(x_{i,\cdot}), q_{n,i}(x_{i,\cdot}) \right]$$

und

$$(16) \quad p_{n,i} \left[x_i, \xi_{n,i}^2(x_{i,\cdot}) \right] = 8d_2^2 \left[p_{n,i}(x_{i,\cdot}), q_{n,i}(x_{i,\cdot}) \right] - 4 q_{n,i}(x_i, N_{n,i}(x_i)).$$

Mithin ist

$$(17) \quad \frac{1}{4} p_{n,i} \left[x_i, \xi_{n,i}^2(x_{i,\cdot}) \right] = - p_{n,i} \left[x_i, \xi_{n,i}(x_{i,\cdot}) \right] - q_{n,i}(x_i, N_{n,i}(x_i))$$

und daher

$$\begin{aligned} & - \frac{1}{4} p_{n,i} \left[x_i, \xi_{n,i}^2(x_{i,\cdot}) \right] = \\ & - \frac{1}{2} p_{n,i} \left[x_i, \xi_{n,i}^2(x_{i,\cdot}) \right] - p_{n,i} \left[x_i, \xi_{n,i}(x_{i,\cdot}) \right] - \\ & q_{n,i}(x_i, N_{n,i}(x_i)). \end{aligned}$$

Dies eingesetzt in (14) führt uns zur stochastischen Entwicklung (3). ■

Die definierenden Eigenschaften eines Lindeberg-Feldes sind u.U. schwierig nachzuweisen. Das folgende Lemma erlaubt es, unter gewissen Umständen diese Eigenschaften für ein anderes Feld von Funktionen zu prüfen.

3 Lemma

Sind

$$\left\{ \xi_{n,i} : 1 \leq i \leq n-1, n > 1 \right\}$$

und

$$\left\{ h_{n,i} : 1 \leq i \leq n-1, n > 1 \right\}$$

Familien von $L^2(P_{n,i,i+1})$ -Funktionen und gilt

$$(1) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\left(\xi_{n,i} - h_{n,i} \right)^2 \right] = 0,$$

so ist eine der beiden Funktionenfamilien genau dann ein Lindeberg-Feld, wenn die andere ein solches ist.

Wenn (1) erfüllt ist und eine der beiden Familien ein Lindeberg-Feld ist, so folgt

$$(2) \quad \lim_{n \rightarrow \infty} P_n \left[\left\{ \sum_{i=1}^{n-1} \left(\xi_{n,i}(x_i, x_{i+1}) - p_{n,i}(x_i, \xi_{n,i}(x_i, \cdot)) \right) - \sum_{i=1}^{n-1} \left(h_{n,i}(x_i, x_{i+1}) - p_{n,i}(x_i, h_{n,i}(x_i, \cdot)) \right) \right\}^2 \right] = 0$$

und

$$(3) \quad \lim_{n \rightarrow \infty} P_n \left[\sum_{i=1}^{n-1} \left| p_{n,i}(x_i, \xi_{n,i}^2(x_i, \cdot)) - p_{n,i}(x_i, h_{n,i}^2(x_i, \cdot)) \right| \right] = 0.$$

Beweis

Angenommen, es gilt (1), und

$$\left\{ \xi_{n,i} : 1 \leq i \leq n-1, n > 1 \right\}$$

ist ein Lindeberg-Feld. Dann ist wegen

$$P_{n,i,i+1} \left[h_{n,i}^2 \right] \leq$$

$$2P_{n,i,i+1} \left[\xi_{n,i}^2 \right] + 2P_{n,i,i+1} \left[\left(h_{n,i} - \xi_{n,i} \right)^2 \right]$$

und (1) sowie 1 (1) für die $\xi_{n,i}$ die Eigenschaft 1 (1) für die $h_{n,i}$ offenbar erfüllt. Ferner gilt

$$(4) \quad P_{n,i,i+1} \left[h_{n,i}^2 \cdot 1_{\left\{ |h_{n,i}| > \varepsilon \right\}} \right] \leq$$

$$2P_{n,i,i+1} \left[\left(h_{n,i} - \xi_{n,i} \right)^2 \cdot 1_{\left\{ |h_{n,i}| > \varepsilon \right\}} \right] +$$

$$2P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |h_{n,i}| > \varepsilon \right\}} \right] \leq$$

$$2P_{n,i,i+1} \left[\left(h_{n,i} - \xi_{n,i} \right)^2 \right] + 2P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |h_{n,i}| > \varepsilon \right\}} \right].$$

Nun ist

$$(6) \quad P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |h_{n,i}| > \varepsilon \right\}} \right] \leq$$

$$P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |\xi_{n,i}| > \frac{\varepsilon}{2} \right\}} \right] +$$

$$P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |h_{n,i} - \xi_{n,i}| > \frac{\varepsilon}{2} \right\}} \right].$$

Der zweite Summand kann wie folgt abgeschätzt werden

$$(7) \quad P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |h_{n,i} - \xi_{n,i}| > \frac{\varepsilon}{2} \right\}} \right] \leq$$

$$P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |\xi_{n,i}| > 1 \right\}} \right] + P_{n,i,i+1} \left[|h_{n,i} - \xi_{n,i}| > \frac{\varepsilon}{2} \right] \leq$$

$$P_{n,i,i+1} \left[\xi_{n,i}^2 \cdot 1_{\left\{ |\xi_{n,i}| > 1 \right\}} \right] + \frac{2}{\varepsilon} P_{n,i,i+1} \left[\left(h_{n,i} - \xi_{n,i} \right)^2 \right].$$

Setzt man die Beziehungen (6) und (7) in (5) ein und summiert über $i = 1, \dots, n-1$, so erhält man nun die Bedingung 1 (2) für die $h_{n,i}$ aus der entsprechenden Bedingung für die $\xi_{n,i}$ und der

Voraussetzung (1).

Gemäss 8.3 und 8.4 ist

$$\begin{aligned}
 P_n \left[\left\{ \sum_{i=1}^{n-1} \left[\xi_{n,i}(x_i, x_{i+1}) - p_{n,i}(x_i, \xi_{n,i}(x_i, \cdot)) \right] - \right. \right. \\
 \left. \left. \sum_{i=1}^{n-1} \left[h_{n,i}(x_i, x_{i+1}) - p_{n,i}(x_i, h_{n,i}(x_i, \cdot)) \right] \right\}^2 \right] = \\
 \sum_{i=1}^{n-1} P_n \left[\left\{ \xi_{n,i}(x_i, x_{i+1}) - p_{n,i}(x_i, \xi_{n,i}(x_i, \cdot)) - \right. \right. \\
 \left. \left. h_{n,i}(x_i, x_{i+1}) + p_{n,i}(x_i, h_{n,i}(x_i, \cdot)) \right\}^2 \right] = \\
 \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\left\{ \xi_{n,i} - h_{n,i} \right\}^2 \right] = \\
 \sum_{i=1}^{n-1} \left\{ P_{n,i} \left[p_{n,i} \left(x_i, \xi_{n,i}(x_i, \cdot) - h_{n,i}(x_i, \cdot) \right) \right] \right\}^2 \leq \\
 \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\left\{ \xi_{n,i} - h_{n,i} \right\}^2 \right] \longrightarrow 0.
 \end{aligned}$$

Damit ist die Konvergenz unter (2) nachgewiesen. Für (3) erhalten wir

$$\begin{aligned}
 P_n \left[\sum_{i=1}^{n-1} \left| p_{n,i}(x_i, \xi_{n,i}^2(x_i, \cdot)) - p_{n,i}(x_i, h_{n,i}^2(x_i, \cdot)) \right| \right] \leq \\
 \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\left| \xi_{n,i}^2 - h_{n,i}^2 \right| \right] = \\
 \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\left| \xi_{n,i} - h_{n,i} \right| \left| \xi_{n,i} + h_{n,i} \right| \right] \leq \\
 \sum_{i=1}^{n-1} \left\{ P_{n,i,i+1} \left[\left\{ \xi_{n,i} - h_{n,i} \right\}^2 \right] P_{n,i,i+1} \left[\left\{ \xi_{n,i} + h_{n,i} \right\}^2 \right] \right\}^{1/2} \leq \\
 \left\{ \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\left\{ \xi_{n,i} - h_{n,i} \right\}^2 \right] \right\}^{1/2} \\
 \left\{ \sum_{i=1}^{n-1} P_{n,i,i+1} \left[\left\{ \xi_{n,i} + h_{n,i} \right\}^2 \right] \right\}^{1/2}.
 \end{aligned}$$

Der erste Faktor des letzten Ausdrucks strebt gemäss (1) gegen 0. Der zweite Faktor ist beschränkt, weil die $\xi_{n,i}$ und die $h_{n,i}$ Lindeberg-Felder sind nach Voraussetzung bzw. laut des ersten Teils unseres Lemmas. ■

Aus Satz 2 und Lemma 3 erhalten wir folgendes Korollar, welches im Zusammenhang der Untersuchungen zur lokalen asymptotischen Normalität eine bedeutende Rolle spielt.

4 Korollar

Wenn

$$\left(h_{n,i} : 1 \leq i \leq n-1, n > 1 \right)$$

ein Lindeberg-Feld ist, welches der Bedingung

$$(1) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} p_{n,i,i+1} \left[\left(2 \left(\frac{dq_{n,i}}{dp_{n,i}} \right)^{1/2} - 2 - h_{n,i} \right)^2 \right] = 0$$

genügt, so ist

$$(3) \quad \frac{dQ_n}{dP_n}(\underline{x}) = \frac{dQ_{n,1}}{dP_{n,1}}(x_1) \exp \left\{ \sum_{i=1}^{n-1} h_{n,i}(x_i, x_{i+1}) - \sum_{i=1}^{n-1} p_{n,i} \left[x_i, h_{n,i}(x_i, \cdot) \right] - \frac{1}{2} \sum_{i=1}^{n-1} p_{n,i} \left[x_i, h_{n,i}^2(x_i, \cdot) \right] - \sum_{i=1}^{n-1} q_{n,i} \left[x_i, N_{n,i}(x_i) \right] + o_P(n^0) \right\}.$$

Beweis

Lemma 3 besagt vermittelt (1), dass die Voraussetzungen des Satzes 1 erfüllt sind. Aus 3 (2) erhalten wir, dass wir

$$\sum_{i=1}^{n-1} \xi_{n,i}(x_i, x_{i+1}) - \sum_{i=1}^{n-1} p_{n,i} \left[x_i, \xi_{n,i}(x_i, \cdot) \right]$$

durch

$$\sum_{i=1}^{n-1} h_{n,i}(x_i, x_{i+1}) = \sum_{i=1}^{n-1} p_{n,i} \left[x_i, h_{n,i}(x_i, \cdot) \right]$$

ersetzen können. Entsprechend erlaubt 3 (3) den Übergang von

$$\sum_{i=1}^{n-1} p_{n,i} \left[x_i, \xi_{n,i}^2(x_i, \cdot) \right]$$

zu

$$\sum_{i=1}^{n-1} p_{n,i} \left[x_i, h_{n,i}^2(x_i, \cdot) \right].$$

Daraus erhalten wir die stochastische Entwicklung (3). ■

8. Einige Aussagen ueber Markoff-Prozesse

Gegeben sei ein polnischer Raum $(E, \mathcal{B}(E))$, auf dem für jedes $i \in \mathbb{N}$ ein Markoff-Kern q_i gegeben sei. Q_1 bezeichne eine Verteilung auf $(E, \mathcal{B}(E))$. Wie in 7 wird durch Q_1 und $\left\{ q_i : i \in \mathbb{N} \right\}$ eindeutig der kanonische Markoff-Prozess $\left(E^{\infty}, \mathcal{B}(E^{\infty}), Q, \left\{ X_i : i \in \mathbb{N} \right\} \right)$ bestimmt. Dabei ist X_i die Projektion auf die i -te Komponente. Es sei

$$Q_n(B_1 \times \dots \times B_n) :=$$

$$\int Q_1(dx_1) 1_{B_1}(x_1) \int q_1(x_1, dx_2) 1_{B_2}(x_2) \dots \int q_{n-1}(x_{n-1}, dx_n) 1_{B_n}(x_n).$$

und

$$Q_{i_1, \dots, i_k}(B_1 \times \dots \times B_k) := Q_n(C_1 \times \dots \times C_n)$$

mit

$$C_l := \begin{cases} B_k, & l = i_k \\ E, & \text{sonst} \end{cases}.$$

Für $l \in \mathbb{N}$ bezeichne

$$\mathcal{V}_l := \sigma(X_1, \dots, X_l)$$

und

$$\mathcal{Z}_l := \sigma(X_m : m \geq l).$$

1 Lemma

Es gilt für jedes $l \in \mathbb{N}$ und alle $V \in \mathcal{V}_l$ sowie $Z \in \mathcal{Z}_l$ Q-f.s:

$$Q[1_{ZV} | X_l] = Q[1_V | X_l] Q[1_Z | X_l].$$

■

2 Lemma

Bezeichnen für $n, k \in \mathbb{N}$

$$f : (E^{n+1}, \mathcal{B}(E^{n+1})) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

und

$$g : (E^{k+1}, \mathcal{B}(E^{k+1})) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

messbare Funktionen derart, dass die nachfolgend gebildeten bedingten Erwartungswerte existieren, und ist

$$1 \leq i + k \leq l \leq m,$$

so gilt

$$\begin{aligned} & Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \cdot g(X_l, X_{l+1}, \dots, X_{l+k}) \mid X_l \right] = \\ & Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \mid X_l \right] \cdot Q \left[g(X_l, X_{l+1}, \dots, X_{l+k}) \mid X_l \right]. \end{aligned}$$

Beweis

$f(X_m, X_{m+1}, \dots, X_{m+n})$ ist $\mathcal{X}_l$ -messbar und $g(X_l, X_{l+1}, \dots, X_{l+k})$ ist $\mathcal{Y}_l$ -messbar. Mit Hilfe des Lemmas 1 und algebraischer Induktion erhält man hieraus

$$\begin{aligned} (1) \quad & Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \mid X_l \right] = \\ & Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \mid X_l, \dots, X_1 \right] \quad Q\text{-f.s.} \end{aligned}$$

Folglich ist Q -f.s.

$$\begin{aligned} & Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \mid X_l \right] \cdot Q \left[g(X_l, X_{l+1}, \dots, X_{l+k}) \mid X_l \right] = \\ & Q \left[g(X_l, X_{l+1}, \dots, X_{l+k}) \cdot Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \mid X_l \right] \mid X_l \right] = \\ & Q \left[g(X_l, X_{l+1}, \dots, X_{l+k}) \cdot \right. \\ & \quad \left. Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \mid X_l, \dots, X_1 \right] \mid X_l \right] = \\ & Q \left[Q \left[g(X_l, X_{l+1}, \dots, X_{l+k}) \cdot f(X_m, X_{m+1}, \dots, X_{m+n}) \right. \right. \\ & \quad \left. \left. \mid X_l, \dots, X_1 \right] \mid X_l \right] = \\ & Q \left[f(X_m, X_{m+1}, \dots, X_{m+n}) \cdot g(X_l, X_{l+1}, \dots, X_{l+k}) \mid X_l \right]. \quad \blacksquare \end{aligned}$$

3 Lemma

Sind für $i \in \mathbb{N}$

$$f_i : (E^2, \mathcal{B}(E^2)) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

messbare Funktionen dergestalt, dass alle nachfolgend gebildeten Erwartungswerte existieren, so gilt

$$\begin{aligned} (1) \quad & Q \left[\left\{ \sum_{i=1}^{n-1} f_i(X_l, X_{l+1}) - Q \left[f_i(X_l, X_{l+1}) \mid X_l \right] \right\}^2 \right] = \\ & \sum_{i=1}^{n-1} Q \left[\left\{ f_i(X_l, X_{l+1}) - Q \left[f_i(X_l, X_{l+1}) \mid X_l \right] \right\}^2 \right]. \end{aligned}$$

Speziell ist also auch

$$(2) \quad Q_n \left[\left\{ \sum_{i=1}^{n-1} f_i(x_i, x_{i+1}) - q_i \left(x_i, f_i(x_i, \cdot) \right) \right\}^2 \right] = \\ \sum_{i=1}^{n-1} Q_{i,i+1} \left[\left\{ f_i(x_i, x_{i+1}) - q_i \left(x_i, f_i(x_i, \cdot) \right) \right\}^2 \right].$$

Beweis

Zur Verkürzung der Schreibweise setzen wir

$$f_i := f_i(X_i, X_{i+1})$$

und

$$Q \left(f_i | X_i \right) := Q \left(f_i(X_i, X_{i+1}) | X_i \right).$$

Dann ist die linke Seite von (1) gleich

$$Q \left[\left\{ \sum_{i=1}^{n-1} f_i - Q \left(f_i | X_i \right) \right\}^2 \right] = \\ \sum_{i=1}^{n-1} \sum_{j=1}^{i-1} Q \left[\left\{ f_i - Q \left(f_i | X_i \right) \right\} \left\{ f_j - Q \left(f_j | X_j \right) \right\} \right] + \\ \sum_{i=1}^{n-1} Q \left[\left\{ f_i - Q \left(f_i | X_i \right) \right\}^2 \right] + \\ \sum_{i=1}^{n-1} \sum_{j=i+1}^{n-1} Q \left[\left\{ f_i - Q \left(f_i | X_i \right) \right\} \left\{ f_j - Q \left(f_j | X_j \right) \right\} \right].$$

Die erste Doppelsumme ist wegen Lemma 2 und $Q(\cdot) = Q(Q(\cdot) | X_i)$ gleich

$$\sum_{i=1}^{n-1} \sum_{j=1}^{i-1} Q \left[Q \left\{ f_i - Q \left(f_i | X_i \right) | X_i \right\} Q \left\{ f_j - Q \left(f_j | X_j \right) | X_i \right\} \right],$$

und diese Doppelsumme verschwindet, weil aufgrund der Linearität des Erwartungswerts

$$Q \left\{ f_i - Q \left(f_i | X_i \right) | X_i \right\} = 0$$

Q-f.s. Ganz analog sieht man, dass die andere Doppelsumme gleich 0 ist. Hieraus erhalten wir die Behauptungen des Lemmas. ■

4 Bemerkung

Ist $\mathcal{F}$ eine Unter- σ -Algebra von $\mathcal{B}(E^{\infty})$ und

$$\xi : E^{\infty} \longrightarrow \mathbb{R}$$

eine messbare Funktion derart, dass alle nachfolgend gebildeten Erwartungswerte existieren, so ist Q-f.s.

$$Q \left[\left\{ \xi - Q \left[\xi \mid \mathcal{F} \right] \right\}^2 \mid \mathcal{F} \right] = Q \left[\xi^2 \mid \mathcal{F} \right] - \left\{ Q \left[\xi \mid \mathcal{F} \right] \right\}^2 \quad \blacksquare$$

Wir spezialisieren uns nun auf den Fall zeitlich homogener Markoff-Prozesse, also $q_i = q$ für alle $i \in \mathbb{N}$. Ist $Q_1 = \varepsilon_x$, $x \in E$, so bezeichne Q^x den aus ε_x und q gewonnenen projektiven Limes der endlich dimensional Randverteilungen des zu ε_x und q gehörigen zeitlich homogenen Markoff-Prozesses. Für eine beliebige Startverteilung Q_1 erhalten wir dann :

$$Q(B) = Q_1(Q^x(B)) \quad , \quad B \in \mathcal{B}(E^{\infty}). \quad \blacksquare$$

5 Lemma

Es gelte für Q_1 und q

$$(1) \quad q(x, \cdot) \ll Q_1 \text{ für alle } x \in E.$$

Ist dann $(f_n : n \in \mathbb{N})$ eine Folge messbarer Funktionen

$$f_n : E^2 \longrightarrow \mathbb{R}$$

mit

$$(3) \quad \lim_n f_n(x, y) = 0 \quad \text{für alle } (x, y) \in E^2$$

und

$$(4) \quad \lim_{i=1}^n f_n(x_i, x_{i+1}) = 0 \quad (Q),$$

so gilt für jede weitere Startverteilung P_1 auf $(E, \mathcal{B}(E))$ ebenfalls

$$(5) \quad \lim_{i=1}^n f_n(x_i, x_{i+1}) = 0 \quad (P)$$

für den aus q und P_1 gebildeten projektiven Limes $P = P_1(Q^x(\cdot))$.

Beweis

Gemäss Bauer 19.4 gilt: Aus jeder Teilfolge $(f_m : m \in M)$ von $(f_n : n \in \mathbb{N})$ kann eine Q-f.s. konvergente Teilfolge $(f_l : l \in L)$ ausgesondert werden, $L \subseteq M \subseteq \mathbb{N}$. Bezeichnet demnach

$$A := \left\{ (x_i : i \in \mathbb{N}) : \lim \sum_{i=1}^l f_i(x_i, x_{i+1}) = 0 \right\},$$

so ist

$$Q(A) = 1.$$

Also erhalten wir

$$(6) \quad Q^1(Q^x(A)) = 1.$$

Notieren wir mit B die Menge

$$B := \left\{ x \in E : Q^x(A) = 1 \right\},$$

so folgt aus (6)

$$Q_1(B) = 1$$

und hieraus wegen (1) für alle $x \in E$

$$(7) \quad q(x, B) = 1.$$

Nun ist für alle $x \in E$

$$(8) \quad Q^y(E \times A) = q(y, Q^x(A)).$$

Wenn nämlich $A = A_1 \times \dots \times A_n \times E^\infty$ eine Zylindermenge in E^∞ ist, so ist ja

$$q(y, Q^x(A)) =$$

$$\int q(y, dx) Q^x(A) =$$

$$\int q(y, dx) \int \varepsilon_x(dx_1) 1_{A_1}(x_1) \int q(x_1, dx_2) 1_{A_2}(x_2) \dots$$

$$\dots \int q(x_{n-1}, dx_n) 1_{A_n}(x_n) =$$

$$\int q(y, dx) 1_{A_1}(x) \int q(x, dx_2) 1_{A_2}(x_2) \dots \int q(x_{n-1}, dx_n) 1_{A_n}(x_n) =$$

$$\int q(y, dx_1) 1_{A_1}(x_1) \int q(x_1, dx_2) 1_{A_2}(x_2) \dots \int q(x_{n-1}, dx_n) 1_{A_n}(x_n) =$$

$$\int \varepsilon_y(dx) \int q(x, dx_1) 1_{A_1}(x_1) \int q(x_1, dx_2) 1_{A_2}(x_2) \dots$$

$$\dots \int q(x_{n-1}, dx_n) 1_{A_n}(x_n) =$$

$$Q^y(E \times A_1 \times \dots \times A_n) = Q^y(E \times A).$$

Mit Hilfe des Eindeutigkeitssatzes für Masse ergibt sich hieraus

(8). Weil $q(x, B) = 1$ gemäss (7), folgt nun mit (8):

$$Q^Y(E \times A) = q(y, 1_B(x)Q^X(A)) = 1.$$

Die letzte Gleichheit ist dabei eine Konsequenz der Definition der Menge B . Aus der Voraussetzung (3) folgern wir $E \times A = A$. Somit ist $Q^Y(A) = 1$ für alle $y \in E$. Darum gilt

$$P(A) = P_1(Q^Y(A)) = 1.$$

Hieraus schliessen wir mit Hilfe des bereits oben zitierten Satzes 19.4 aus Bauer die Behauptung. ■

9. Lebesgue-Zerlegung fuer Kerne

Betrachte zwei messbare Räume $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$. Ein Kern von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$ ist eine Abbildung

$$q : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}$$

derart, dass

$$(1) \quad q(\omega_1, \cdot) : \Omega_2 \longrightarrow \mathbb{R}$$

ein Mass auf $(\Omega_2, \mathcal{A}_2)$ ist für jedes feste $\omega_1 \in \Omega_1$ und

$$(2) \quad q(\cdot, A_2) : \Omega_1 \longrightarrow \mathbb{R}$$

eine $\mathcal{A}_1$ - $\mathcal{B}(\mathbb{R})$ -messbare Abbildung ist für jedes $A_2 \in \mathcal{A}_2$.

1 Definition

Sind q und p zwei Kerne von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$, so heisst eine $\mathcal{A}_1 \otimes \mathcal{A}_2$ - $\mathcal{B}(\mathbb{R})$ messbare Funktion

$$f : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}_+$$

eine Dichte von q bzgl. p , falls für jedes $\omega_1 \in \Omega_1$ und alle $A_2 \in \mathcal{A}_2$ gilt

$$q(\omega_1, A_2) = p(\omega_1, 1_{A_2}(\cdot) f(\omega_1, \cdot)) =$$

$$\int p(\omega_1, d\omega_2) 1_{A_2}(\omega_2) f(\omega_1, \omega_2).$$

Für f schreiben wir auch $\frac{dq}{dp}$. ■

2 Bemerkung

Ist $g : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}$ eine $\mathcal{A}_1 \otimes \mathcal{A}_2$ -messbare Funktion und $q(\omega_1, g(\omega_1, \cdot))$ erklärt für alle $\omega_1 \in \Omega_1$, so ist

$$\omega_1 \longrightarrow q(\omega_1, g(\omega_1, \cdot))$$

eine $\mathcal{A}$ - $\mathcal{B}(\mathbb{R})$ -messbare Funktion (vgl. Gänssler/Stute, 1.8.7). Ist f eine Dichte von q bzgl. p , so gilt

$$q(\omega_1, g(\omega_1, \cdot)) = p(\omega_1, f(\omega_1, \cdot) g(\omega_1, \cdot)).$$
 ■

3 Definition

p und q seien Kerne von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$ und

$$f : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}_+$$

eine $\mathcal{A}_1 \otimes \mathcal{A}_2$ - $\mathcal{B}(\mathbb{R})$ -messbare Funktion. Des weiteren sei

$$\left\{ N(\omega_1) : \omega_1 \in \Omega_1 \right\}$$

eine Familie von messbaren Mengen aus $\mathcal{A}_2$ mit

$$p(\omega_1, N(\omega_1)) = 0$$

für alle $\omega_1 \in \Omega_1$. Wenn dann

$$q(\omega_1, A_2) = p\left(\omega_1, 1_{A_2}(\cdot) f(\omega_1, \cdot)\right) + q(\omega_1, A_2 \cap N(\omega_1))$$

gilt, so heisst $\left[f, \left\{ N(\omega_1) : \omega_1 \in \Omega_1 \right\} \right]$ eine Lebesgue-Zerlegung

von q bzgl. p . Wir schreiben für f auch $\frac{dq}{dp}$. ■

4 Bemerkung

Offenbar ist der Fall der Existenz einer Dichte von q bzgl. p ein Spezialfall der Lebesgue-Zerlegung. ■

Wir fragen nun nach der Existenz und der Eindeutigkeit von Lebesgue-Zerlegungen.

5 Satz

Sind $\left[f, \left\{ N(\omega_1) : \omega_1 \in \Omega_1 \right\} \right]$ und $\left[g, \left\{ M(\omega_1) : \omega_1 \in \Omega_1 \right\} \right]$

zwei Lebesgue-Zerlegungen von q bzgl. p , so gilt

$$(1) \quad f(\omega_1, \cdot) = g(\omega_1, \cdot) \quad p(\omega_1, \cdot)\text{-f.s. für alle } \omega_1 \in \Omega_1.$$

$$(2) \quad q(\omega_1, N(\omega_1) \Delta M(\omega_1)) = 0 \quad \text{für alle } \omega_1 \in \Omega_1.$$

Dabei bezeichnet für beliebige Mengen A und B $A \Delta B$ die symmetrische Differenz dieser Mengen.

Beweis

Bezeichne $\bar{A}$ das mengentheoretische Komplement einer Menge A .

Für jedes $\omega_1 \in \Omega_1$ und $A_2 \in \mathcal{A}_2$ gilt:

$$p\left(\omega_1, f(\omega_1, \cdot) 1_{A_2 \cap \overline{N(\omega_1)}}(\cdot)\right) =$$

$$q\left(\omega_1, 1_{A_2 \cap \overline{N(\omega_1)}}(\cdot) \overline{M(\omega_1)}(\cdot)\right) =$$

$$p\left(\omega_1, g(\omega_1, \cdot) 1_{A_2 \cap \overline{N(\omega_1)}}(\cdot)\right).$$

Mithin ist

$$g(\omega_1, \cdot) = f(\omega_1, \cdot) \quad \text{auf } \overline{N(\omega_1)} \setminus \overline{M(\omega_1)} \quad p(\omega_1, \cdot)\text{-f.s.}$$

für alle $\omega_1 \in \Omega_1$. Nun ist aber für jedes $\omega_1 \in \Omega_1$

$$p\left(\omega_1, \overline{N(\omega_1)} \setminus \overline{M(\omega_1)}\right) = 1.$$

Folglich ist

$$g(\omega_1, \cdot) = f(\omega_1, \cdot) \quad p(\omega_1, \cdot)\text{-f.s.}$$

für alle $\omega_1 \in \Omega_1$. Dies beweist (1).

Um (2) einzusehen, bemerken wir, dass

$$\begin{aligned} (3) \quad q\left(\omega_1, \overline{N(\omega_1)} \setminus \overline{M(\omega_1)}\right) &= \\ q\left(\omega_1, \overline{M(\omega_1)}\right) - p\left(\omega_1, f(\omega_1, \cdot) 1_{\overline{M(\omega_1)}}(\cdot)\right) &= \\ q\left(\omega_1, \overline{M(\omega_1)}\right) - p\left(\omega_1, g(\omega_1, \cdot) 1_{\overline{M(\omega_1)}}(\cdot)\right) \end{aligned}$$

wegen (1). Der letzte Ausdruck ist gleich 0. Ganz entsprechend berechnet man

$$(4) \quad q\left(\omega_1, \overline{N(\omega_1)} \cap \overline{M(\omega_1)}\right).$$

(3) und (4) zusammen ergeben (2). ■

6 Satz

Sind p , q und r Kerne von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$ und besitzen p und q Dichten $\frac{dp}{dr}$ und $\frac{dq}{dr}$ bzgl. r , so existiert eine Lebesgue-Zerlegung von q bzgl. p , nämlich

$$\frac{dq}{dp}(\omega_1, \omega_2) := \begin{cases} \frac{dq}{dr}(\omega_1, \omega_2) \left(\frac{dp}{dr}(\omega_1, \omega_2) \right)^{-1}, & \text{falls } \frac{dp}{dr}(\omega_1, \omega_2) \neq 0 \\ 0, & \text{sonst} \end{cases}$$

und

$$N(\omega_1) := \left\{ \omega_2 \in \Omega_2 : \frac{dp}{dr}(\omega_1, \omega_2) = 0 \right\}.$$

Beweis

Nachrechnen ■

7 Bemerkung

Sei $g: \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}$ eine $\mathcal{A}_1 \otimes \mathcal{A}_2$ -messbare Funktion und $q(\omega_1, g(\omega_1, \cdot))$

erklärt für alle $\omega_1 \in \Omega_1$. Ist dann $\left[\frac{dq}{dp}, \left\{ N(\omega_1) : \omega_1 \in \Omega_1 \right\} \right]$ eine Lebesgue-Zerlegung von q bzgl. p , so gilt

$$q(\omega_1, g(\omega_1, \cdot)) = p\left[\omega_1, \frac{dq}{dp}(\omega_1, \cdot) g(\omega_1, \cdot)\right] + q\left[\omega_1, 1_{N(\omega_1)}(\cdot) g(\omega_1, \cdot)\right].$$

Beweis

Die Aussage ergibt sich aus der Definition der Lebesgue-Zerlegung mittels algebraischer Induktion (vgl. Gänssler/Stute). ■

8 Bemerkung

Sind $(\Omega_i, \mathcal{A}_i)$, $1 \leq i \leq n+1$ messbare Räume und q_i Kerne von $(\Omega_i, \mathcal{A}_i)$ nach $(\Omega_{i+1}, \mathcal{A}_{i+1})$, $1 \leq i \leq n$, so ist der Produktkern

$$\bigotimes_{i=1}^n q_i : \Omega_1 \times \left[\bigotimes_{j=2}^{n+1} \mathcal{A}_j \right] \longrightarrow \mathbb{R}_+$$

definiert durch die Eigenschaft

$$\begin{aligned} \bigotimes_{i=1}^n q_i(\omega_1, A_2 \times \dots \times A_n) &= \int q_1(\omega_1, d\omega_2) 1_{A_2}(\omega_2) \int q_2(\omega_2, d\omega_3) 1_{A_3}(\omega_3) \\ &\quad \dots \int q_n(\omega_n, d\omega_{n+1}) 1_{A_{n+1}}(\omega_{n+1}) \end{aligned}$$

(vgl. Gänssler/Stute, 1.9). Für die Dichte von Produktkernen gilt:

9 Lemma

Mit den Bezeichnungen von 8 mögen Dichten $\frac{dq_i}{dp_i}$ von q_i bzgl. p_i

existieren. Dann besitzt $\bigotimes_{i=1}^n q_i$ eine Dichte bzgl. $\bigotimes_{i=1}^n p_i$, nämlich

$$\frac{d \bigotimes_{i=1}^n q_i}{d \bigotimes_{i=1}^n p_i}(\omega_1, \dots, \omega_{n+1}) = \prod_{i=1}^n \frac{dq_i}{dp_i}(\omega_i, \omega_{i+1}).$$

Hinsichtlich der Lebesgue-Zerlegung von Produkt-Kernen erhalten wir folgendes Lemma.

10 Lemma

Sind q_i , p_i und r_i Kerne von $(\Omega_i, \mathcal{A}_i)$ nach $(\Omega_{i+1}, \mathcal{A}_{i+1})$, $1 \leq i \leq n$, und existieren Dichten von q_i bzgl. r_i und p_i bzgl. r_i , so existiert die Lebesgue-Zerlegung von $\bigotimes_{i=1}^n q_i$ bzgl. $\bigotimes_{i=1}^n p_i$.

Eine solche ist gegeben durch

$$\left\{ \frac{d \bigotimes_{i=1}^n q_i}{d \bigotimes_{i=1}^n p_i}, \left\{ N(\omega_1) : \omega_1 \in \Omega_1 \right\} \right\},$$

wenn

$$N(\omega_1) := \left\{ (\omega_2, \dots, \omega_{n+1}) : \prod_{i=1}^n \frac{dq_i}{dr_i} (\omega_i, \omega_{i+1}) = 0 \right\}$$

und

$$\frac{d \bigotimes_{i=1}^n q_i}{d \bigotimes_{i=1}^n p_i} (\omega_1, \dots, \omega_{n+1}) := \begin{cases} \left[\prod_{i=1}^n \frac{dq_i}{dr_i} (\omega_i, \omega_{i+1}) \right] \left[\prod_{i=1}^n \frac{dp_i}{dr_i} (\omega_i, \omega_{i+1}) \right]^{-1}, & \text{für } (\omega_2, \dots, \omega_{n+1}) \notin N(\omega_1) \\ 0, & \text{sonst} \end{cases}$$

Beweis

Die Aussage folgt problemlos aus 6 und 9. ■

11 Bemerkung

Sind P und Q zwei Verteilungen auf $(\Omega_1, \mathcal{A}_1)$ und p und q zwei Kerne

von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$ derart, dass Dichten $\frac{dp}{dr}$ und $\frac{dq}{dr}$ bzgl.

eines weiteren Kernes r existieren, so ist eine Lebesgue-Zerlegung

von $Q \otimes q$ bzgl. $P \otimes p$ gegeben durch $\left(\frac{dQ \otimes q}{dP \otimes p}, N \right)$, wobei

$$N := \left\{ (\omega_1, \omega_2) : \frac{dP}{d(Q+P)} (\omega_1) \frac{dp}{dr} (\omega_1, \omega_2) = 0 \right\}$$

und

$$\frac{dQ \otimes q}{dP \otimes p} (\omega_1, \omega_2) := \frac{dQ}{dP} (\omega_1) \frac{dq}{dp} (\omega_1, \omega_2).$$
■

10. Differenzierbare Mannigfaltigkeiten

Wir führen in diesem Abschnitt den Begriff einer differenzierbaren Mannigfaltigkeit ein, wie wir ihn in dieser Arbeit verwenden. Die im folgenden vorgestellte Konzeption weicht von der üblicherweise in der Literatur anzutreffenden ab. Sie folgt dem Beispiel, das in Spallek, Kurven und Karten, für differenzierbare Mannigfaltigkeiten im $\mathbb{R}^n$ vorgezeichnet ist. Wir wählen den unten dargestellten Begriff der differenzierbaren Mannigfaltigkeit insbesondere deshalb, weil der mit ihm verbundene Begriff des Tangentialraums anschaulich ist und gerade unseren Zwecken angemessen erscheint.

In diesem Abschnitt bezeichnen, sofern nicht anders vermerkt, $(B, |\cdot|)$, $(B', |\cdot|')$ und $(B'', |\cdot|'')$ stets Banach-Räume über $\mathbb{R}$, entsprechend A , A' , A'' Teilmengen von B bzw. B' bzw. B'' .

1 Definition

Eine Abbildung $f : A \longrightarrow A'$ heisst (stetig) differenzierbar, falls für jedes $a \in A$ eine offene Umgebung $U \subset B$ mit $a \in U$ und eine stetig differenzierbare Abbildung $F : U \longrightarrow B'$ existieren, so dass

$$F|_{U \cap A} = f|_{U \cap A}.$$

f heisst Diffeomorphismus, wenn f bijektiv und sowohl f als auch f^{-1} stetig differenzierbar sind. Wir erinnern daran, dass F in x' differenzierbar ist, falls ein beschränkter linearer Operator $DF_{x'} : B \longrightarrow B'$ existiert mit der Eigenschaft

$$|F(x) - F(x') - DF_{x'}(x - x')| = o(|x - x'|).$$

2 Definition

Sei $a \in A \subset B$.

(1) Die Menge

$Tg_a A := \{x \in B : \text{für jede in einer offenen Umgebung } U \subseteq B \text{ von } a \text{ definierte differenzierbare Funktion}$

$F : U \longrightarrow \mathbb{R}, \text{ die auf } A \text{ verschwindet, gilt}$

$$DF_a(x) = 0\}$$

heisst Tangentialraum von A in a .

(2) Als geometrischen Tangentialraum von A in a definieren wir

$$T_{\mathcal{E}_a}^g A := \{ x \in B : \text{es gibt Folgen } (c_n : n \in \mathbb{N}) \subset \mathbb{R} \text{ und} \\ (a_n : n \in \mathbb{N}) \subset A \text{ mit } |a_n - a| \longrightarrow 0 \text{ und} \\ |x - c_n(a_n - a)| \longrightarrow 0 \}.$$

3 Bemerkung

(1) Ist $A \subseteq B$ offen, so ist für $a \in A$ $T_{\mathcal{E}_a} A = T_{\mathcal{E}_a}^g A = B$.

Sei A eine beliebige Teilmenge von B . Dann sind folgende Aussagen wahr.

(2) $T_{\mathcal{E}_a}^g A \subseteq T_{\mathcal{E}_a} A$ für alle $a \in A$.

(3) Ist $U \subseteq B$ offen und sind $F, G : U \longrightarrow B'$ zwei differenzierbare Abbildungen mit $F|_{U \cap A} = G|_{U \cap A}$, so gilt für jedes $a \in A$

$$DF_a|_{T_{\mathcal{E}_a} A} = DG_a|_{T_{\mathcal{E}_a} A}.$$

(4) Ist $U \subseteq B$ offen, $F : U \longrightarrow B'$ differenzierbar, $a \in A \cap U$, $a' := F(a)$ sowie $F(A \cap U) \subseteq A'$, so ist

$$DF_a(T_{\mathcal{E}_a} A) \subseteq T_{\mathcal{E}_a} A'.$$

4 Definition

$f : A \longrightarrow A'$ sei differenzierbar. Wegen 3 (3) ist folgende Definition eindeutig. Wenn $a \in A$ und $U \subseteq B$ irgendeine offene Umgebung von a ist sowie $F : U \longrightarrow B'$ irgendeine differenzierbare Funktion mit $F|_{U \cap A} = f|_{U \cap A}$, dann setze für $a' := f(a)$

$$Df_a := DF_a|_{T_{\mathcal{E}_a} A} : T_{\mathcal{E}_a} A \longrightarrow T_{\mathcal{E}_a} A'$$

Df_a heisst Ableitung von f in a .

5 Lemma

(1) Sind $f : A \longrightarrow A'$ und $g : A' \longrightarrow A''$ differenzierbar, so auch $g \circ f$ und es gilt

$$D(g \circ f)_a = Dg_{f(a)} \circ Df_a.$$

(2) Falls $f : A \longrightarrow A'$ ein Diffeomorphismus ist, so ist Df_a bijektiv und für alle $a \in A$ ist

$$(Df_a)^{-1} = D(f^{-1})_{f(a)}.$$

6 Definition

Eine Teilmenge $M \subseteq B$ heisst differenzierbare B' -Mannigfaltigkeit,

wenn zu jedem $m \in M$ eine relativ offene Umgebung $U \subseteq M$ und eine offene Teilmenge $V' \subseteq \mathbb{B}'$ sowie ein Diffeomorphismus $\Psi : U \longrightarrow V'$ existiert. Das Tripel (U, Ψ, V') heisst eine Karte von M . Ein Atlas von M ist eine Familie $\left\{ (U_i, \Psi_i, V'_i) : i \in I \right\}$ von Karten derart, dass $\bigcup_i U_i = M$.

Ist speziell $\mathbb{B}' = \mathbb{H}$ ein Hilbertraum, so sprechen wir auch von einer differenzierbaren Hilbert-Mannigfaltigkeit, ist $\mathbb{B}'$ ein euklidischer Raum entsprechend von einer euklidischen differenzierbaren Mannigfaltigkeit. Die Dimension einer differenzierbaren Mannigfaltigkeit ist die Dimension von $\mathbb{B}'$. ■

7 Bemerkung

Unsere Definition einer differenzierbaren Mannigfaltigkeit ist ein Spezialfall der üblicherweise in der Literatur zu findenden Begriffsbildungen. Genauer gilt:

Ist M eine differenzierbare Mannigfaltigkeit im Sinne der Definition 6, so auch im Sinne der Definition aus Choquet-Bruhat e.a., VII. ■

8 Lemma

Für jede differenzierbare $\mathbb{B}'$ -Mannigfaltigkeit $M \subseteq \mathbb{B}$ und $m \in M$ sowie jede Karte (U, Ψ, V') mit $m \in U$ gilt

$$T_m M = T_m^g M = \left(D\Psi \right)^{-1} (V').$$

9 Lemma

Für jede differenzierbare $\mathbb{B}'$ -Mannigfaltigkeit M und $m \in M$ sowie eine beliebig vorgegebene Folge reeller Zahlen δ_n mit $\delta_n \longrightarrow 0$ gilt

$$T_m M = \{ x \in \mathbb{B} : \text{es gibt eine Folge von Elementen } m_n \in M \text{ mit } m_n \longrightarrow m \text{ und } |x - \delta_n (m_n - m)| \longrightarrow 0 \}.$$

Beweis

Betrachte eine Karte (U, Ψ, V') von M mit $m \in U$ und ein $x \in T_m M$. Dann ist $x' := D\Psi_m(x) \in \mathbb{B}'$ und deshalb $\Psi(m) + x' \delta_n^{-1} \in V'$ für hinreichend grosse $n \in \mathbb{N}$. Definiere $m_n := \Psi^{-1}(\Psi(m) + x' \delta_n^{-1})$. Dann gilt

$$|\delta_n(m_n - m) - x| =$$

$$|\delta_n \left[\Psi^{-1}(\Psi(m) + x' \delta_n^{-1}) - \Psi^{-1} \Psi(m) \right] - \left[D\Psi_m \right]^{-1}(x')| \leq$$

$$|\delta_n \left[D\Psi_m \right]^{-1} \left[x' \delta_n^{-1} \right] - \left[D\Psi_m \right]^{-1}(x')| \longrightarrow 0.$$

■

10 Lemma

M sei eine differenzierbare Hilbert-Mannigfaltigkeit in einem Banachraum. Für eine stetig differenzierbare Abbildung $f : M \longrightarrow \mathbb{R}^k$ und ein $c \in \mathbb{R}^k$ definiere die Menge

$$M' := \{ m \in M : f(m) = c \}.$$

Falls für jedes $m \in M'$ Df_m surjektiv ist, dann ist M' eine differenzierbare Hilbert-Mannigfaltigkeit.

Für ihren

Tangentialraum gilt

$$T_{\mathbb{E}_m} M' = \{ t : Df_m(t) = 0 \}.$$

Beweis

Wir geben eine Karte von M' an für beliebige $m_0 \in M'$. Ohne Einschränkung sei $c = 0 \in \mathbb{R}^k$. Zunächst existiert eine Karte (U, Ψ, V) von M mit $m_0 \in U$. Mit Φ bezeichnen wir den Umkehrdiffeomorphismus zu Ψ , $h_0 := \Psi(m_0)$. $\mathbb{H}$ sei der zugrundeliegende Hilbertraum, in dem M modelliert wird. Die Abbildung $f \circ \Phi : V \longrightarrow \mathbb{R}^k$ ist stetig differenzierbar, und ihre Ableitung ist

$$Df_{m_0} \circ D\Phi_{h_0} = D(f \circ \Phi)_{h_0} : \mathbb{H} \longrightarrow \mathbb{R}^k.$$

Sie ist surjektiv. Sei

$$K := \left\{ x \in \mathbb{H} : D(f \circ \Phi)_{h_0}(x) = 0 \right\}$$

der Kern dieser Ableitung. Dann existieren k Vektoren $e_1, \dots, e_k \in \mathbb{H}$, die orthogonal zu K und normiert sind. Definiere eine Abbildung

$$F : V \longrightarrow \mathbb{H}$$

$$F(x) := \sum_{i=1}^k [(f \circ \Phi)(x)]_i e_i + \pi_K(x),$$

wobei $[(f \circ \Phi)(x)]_i$ die i -te Komponente von $(f \circ \Phi)(x)$ bezeichnet und π_K die orthogonale Projektion auf K . F ist stetig differenzierbar mit Ableitung

$$DF_{h_0}(x) = \sum_{i=1}^k [D(f \circ \Phi)_{h_0}(x)]_i e_i + \pi_K(x).$$

Ausserdem ist DF_{h_0} bijektiv und beschränkt. Gemäss Dieudonné,

10.25 existiert dann eine Umgebung $\hat{V}$ von h_0 und $\hat{W}$ von $F(h_0)$ derart, dass $F: \hat{V} \rightarrow \hat{W}$ ein Diffeomorphismus ist. K ist als abgeschlossener linearer Teilraum des Hilbert-Raums H selbst ein Hilbert-Raum, und $\bar{V} := \hat{W} \cap K$ ist eine offene Teilmenge von K .

$$\bar{\Psi} := F^{-1}|_{\bar{V}}: \bar{V} \rightarrow \bar{U} := F^{-1}(\hat{W}) \cap F^{-1}(K)$$

ist ein Diffeomorphismus. Weil

$$F^{-1}(K) = \{ h \in V : F(h) \in K \} =$$

$$\left\{ [(f \circ \Phi)(h)]_i = 0 \text{ für alle } i = 1, \dots, k \right\},$$

ist $\bar{U}$ offene Teilmenge von $\{ h \in V : (f \circ \Phi)(h) = 0 \}$. Also ist $(\bar{U}, \bar{\Psi}, \bar{V})$ eine Karte von $\{ h \in V : (f \circ \Phi)(h) = 0 \}$ mit $h_0 \in \bar{U}$.

Definiere nun

$$U' := \Phi(U)$$

$$\Psi' := \bar{\Psi} \circ \Psi|_{U'}$$

$$V' := \bar{V}.$$

Dann ist (U', Ψ', V') die gesuchte Karte von M' mit $m_0 = \Phi(h_0) \in U'$.

Die Aussage die Tangentialräume betreffend erhält man sogleich aus

$$T_{m_0} M' = \left[D\Psi'_{m_0} \right]^{-1}(K).$$

Es ist aber $D\Psi'_{m_0}(t) \in K$ dann und nur dann, wenn $Df_{m_0}(t) = 0$. ■

11. Beweistechnische Hilfsmittel

In diesem Abschnitt stellen wir zum Zweck der einfachen Referenz und um lästige Rechnereien im Haupttext zu vermeiden einige Aussagen zusammen.

1 Totalvariation

$(\Omega, \mathcal{A})$ bezeichne einen messbaren Raum und P und Q zwei Verteilungen darauf.

(1) Die Totalvariation von P und Q ist definiert durch

$$d_1(P, Q) := \sup \{ |P(A) - Q(A)| : A \in \mathcal{A} \}.$$

(2) Ist $\left[\frac{dQ}{dP}, N \right]$ die Lebesgue-Zerlegung von Q bzgl. P , so gilt

$$d_1(P, Q) = \frac{1}{2} P \left[\left| \frac{dQ}{dP} - 1 \right| \right] + \frac{1}{2} Q(N)$$

(vgl. Strasser (1), 2.2).

(3) d_1 ist eine Metrik auf der Menge aller Verteilungen auf $(\Omega, \mathcal{A})$

(vgl. Strasser (1), 2.5).

(4) Sind P und Q dominiert von einem σ -endlichen Mass R , so ist

$$d_1(P, Q) = \frac{1}{2} R \left[\left| \frac{dQ}{dR} - \frac{dP}{dR} \right| \right]$$

(vgl. Strasser (1), 2.4).

2 Hellinger-Abstand

P und Q seien zwei Verteilungen auf dem messbaren Raum $(\Omega, \mathcal{A})$, die von einem σ -endlichen Mass R dominiert werden.

(1) Der Hellinger-Abstand von P und Q ist erklärt durch

$$d_2(P, Q) := \left\{ \frac{1}{2} R \left[\left(\frac{dP}{dR} \right)^{1/2} - \left(\frac{dQ}{dR} \right)^{1/2} \right]^2 \right\}^{1/2}.$$

(2) Die Grösse

$$\alpha(P, Q) := R \left[\left(\frac{dP}{dR} \right)^{1/2} \left(\frac{dQ}{dR} \right)^{1/2} \right]$$

heisst Affinität der Masse P und Q .

Es gelten folgende Beziehungen.

$$(3) \quad d_2^2(P, Q) = 1 - \alpha(P, Q)$$

(vgl. Strasser (1), 2.9).

$$(4) \quad \alpha(P, Q) = P \left[\left(\frac{dQ}{dP} \right)^{1/2} \right]$$

für die Lebesgue-Zerlegung $\left[\left(\frac{dQ}{dP} \right), N \right]$ von Q bzgl. P

(vgl. Strasser (1), 2.10).

$$(5) \quad P \left[2 \left(\frac{dQ}{dP} \right)^{1/2} - 2 \right] = -2d_2^2(P, Q)$$

(vgl. Strasser (1), 2.13).

$$(6) \quad P \left[\left\{ 2 \left(\frac{dQ}{dP} \right)^{1/2} - 2 \right\}^2 \right] = 8d_2^2(P, Q) - 4Q(N)$$

(vgl. Strasser (1), 2.12 und 2.13).

$$(7) \quad d_2^2(P, Q) \leq d_1(P, Q) \leq \left[d_2^2(P, Q) \left(2 - d_2^2(P, Q) \right) \right]^{1/2}$$

(vgl. Strasser (1), 2.15).

(8) d_2 ist eine Metrik auf der Menge der von R dominierten Verteilungen (vgl. Strasser (1), 2.14).

(9) Sind $(\Omega_i, \mathcal{A}_i)$, $1 \leq i \leq n \in \mathbb{N}$ messbare Räume und P_i und Q_i Verteilungen auf $(\Omega_i, \mathcal{A}_i)$, die von einem σ -endlichen Mass R_i dominiert werden, so ist

$$\alpha \left(\bigotimes_{i=1}^n P_i, \bigotimes_{i=1}^n Q_i \right) = \prod_{i=1}^n \alpha(P_i, Q_i).$$

(vgl. Strasser (1), 2.16).

4 Taylor-Entwicklung

(1) Für alle $x > -2$ gilt

$$2 \log \left(\frac{x}{2} + 1 \right) = x - \frac{x^2}{4} + \frac{x^2}{4} (1 - r(x))$$

mit

$$r(x) = \int_0^1 \frac{2(1-s)}{\left(1 + \frac{sx}{2}\right)^2} ds.$$

(2) Die Funktion r aus (1) ist auf dem offenen Intervall $(-1, 1)$ gleichmässig stetig. Es ist

$$|r(x) - r(y)| \leq c|x - y|$$

für alle $x, y \in (-1, 1)$ und ein $c \in \mathbb{R}$.

Beweis

Die Taylor-Entwicklung um $x = 0$ von $\log \left(\frac{x}{2} + 1 \right)$ bis zum linearen Glied ergibt

$$2 \log \left[\frac{x}{2} + 1 \right] = x - 2 \int_0^x (x-t)(t+2)^{-2} dt.$$

Nun ist

$$2 \int_0^x (x-t)(t+2)^{-2} dt =$$

$$\frac{x^2}{4} \int_0^1 \frac{2(1-s)}{\left[1 + \frac{sx}{2}\right]^2} ds.$$

Daraus folgt (1). Es ist

$$|r(x) - r(y)| \leq$$

$$\left| \int_0^1 \frac{8(1-s)}{(2+sx)^2} - \frac{8(1-s)}{(2+sy)^2} ds \right| \leq$$

$$8 \int_0^1 | (2+sy)^2 - (2+sx)^2 | ds$$

für $|x|, |y| < 1$, weil dann $(2+sx)^2 \geq 1$ für alle $s \in [0,1]$.

Damit schliessen wir weiter:

$$8 \int_0^1 | (2+sy)^2 - (2+sx)^2 | ds \leq$$

$$8 \int_0^1 | 2+sy + 2+sx | |s| |x-y| ds \leq$$

$$48 |x-y|.$$

■

5 Stochastische Beschränktheit

Gegeben seien für jedes $n \in \mathbb{N}$ ein Wahrscheinlichkeitsraum $(\Omega_n, \mathcal{A}_n, P_n)$ und Zufallsvariablen

$$X_n, Y_n : (\Omega_n, \mathcal{A}_n) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

(1) $(X_n : n \in \mathbb{N})$ heisst nach P_n -Wahrscheinlichkeit (oder P_n -stochastisch) beschränkt, falls für alle $\varepsilon > 0$ eine Konstante $C(\varepsilon)$ existiert, so dass

$$\limsup_{n \in \mathbb{N}} P_n \{ |X_n| > C(\varepsilon) \} < \varepsilon.$$

(2) Ist $\langle Y_n : n \in \mathbb{N} \rangle$ P_n -stochastisch konvergent gegen 0 und $\langle X_n : n \in \mathbb{N} \rangle$ P_n -stochastisch beschränkt, so strebt die Folge der $X_n Y_n$ in P_n -Wahrscheinlichkeit gegen 0.

Beweis

$$\begin{aligned} P_n \{ |X_n Y_n| > \alpha \} &\leq \\ P_n \{ (|X_n Y_n| > \alpha) \cap (|X_n| > C) \} &+ \\ P_n \{ (|X_n Y_n| > \alpha) \cap (|X_n| \leq C) \} &\leq \\ P_n \{ |X_n| > C \} + P_n \{ |Y_n| > \alpha C^{-1} \}. \end{aligned}$$

Zu jedem $\varepsilon > 0$ existiert ein $C(\varepsilon)$ derart, dass

$$\limsup_{n \in \mathbb{N}} P_n \{ |X_n| > C(\varepsilon) \} < \varepsilon.$$

Daher ist

$$\begin{aligned} \limsup_{n \in \mathbb{N}} P_n \{ |X_n Y_n| > \alpha \} &\leq \\ \limsup_{n \in \mathbb{N}} P_n \{ |X_n| > C(\varepsilon) \} + \limsup_{n \in \mathbb{N}} P_n \{ |Y_n| > \alpha C(\varepsilon)^{-1} \} &< \varepsilon. \end{aligned}$$

Weil $\varepsilon > 0$ beliebig war folgt nun die Behauptung. ■

(3) Gilt für $\langle X_n : n \in \mathbb{N} \rangle$ und ein $q \geq 1$

$$\limsup_{n \in \mathbb{N}} P_n \left[|X_n|^q \right] < \infty,$$

so ist $\langle X_n : n \in \mathbb{N} \rangle$ nach P_n -Wahrscheinlichkeit beschränkt.

Beweis

Folgt sogleich mit Hilfe der Tschebyscheff-Markoff-Ungleichung. ■

6 Lemma

Sei $f : (-1, 1) \rightarrow \mathbb{R}$ eine Funktion und t_n eine Folge in $(0, 1)$, für die $t_n^{-1} - t_{n-1}^{-1}$ beschränkt ist. Genau dann gilt

$$f(t) \rightarrow 0 \text{ für } t \rightarrow 0,$$

wenn

$$f(t_n u_n) \rightarrow 0 \text{ für jede Folge } u_n \text{ mit Grenzwert } +1 \text{ oder } -1.$$

Speziell kann $t_n := n^{-1/2}$ gewählt werden.

Beweis

Folgt aus 1.2.10 aus Pfanzagl (2) durch Betrachtung des links- und rechtsseitigen Grenzwertes. ■

Literatur

Basawa, J.V./ Prakasa Rao, B.L.S., Statistical Inference for Stochastic Processes, Academic Press, London, 1980

Bauer, H., Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie, de Gruyter, Berlin, 1978

Billingsley (1), P., Statistical Inference for Markov Processes, Institute of Mathematical Statistics - University of Chicago Statistical Research Monographs, University of Chicago Press, 1961

Billingsley (2), P., The Lindeberg-Lévy Theorem for Martingales, Proc. Amer. Math. Soc. 12, 788 - 792, 1961

Choquet-Bruhat, Y., e.a., Analysis, Manifolds and Physics, North-Holland, 1977

Dieudonne, J., Grundzüge der modernen Analysis, Bd. 1, 2. Aufl., Vieweg, Braunschweig, 1971

Doob, J.L., Stochastic Processes, Wiley, New York, 1953

Gänssler, P./ Stute, W., Wahrscheinlichkeitstheorie, Springer Verlag, Berlin, 1977

Ibragimov, I.A./ Has'minskii, R.Z., Statistical Estimation, Asymptotic Theory, Springer Verlag, New York, 1981

Kolmogoroff, A.N., Über die Summen durch den Zufall bestimmter unabhängiger Grössen, Math. Ann. 99, 309-319, 1928; 102, 484-488, 1929

Loève, M., Probability Theory I, 4.ed., Springer Verlag, New York, 1977

Le Cam, L., Locally Asymptotically Normal Families, University of California Publ. in Statistics 3, 1960

Le Cam, L., Asymptotic Methods in Statistical Decision Theory,
Springer Verlag, New York, 1986

Pfanzagl (1), J., Contributions to a General Asymptotic
Statistical Theory, Lecture Notes in Statistics 13, Springer
Verlag, New York, 1982

Pfanzagl (2), J., Asymptotic Expansions for General Statistical
Modells, Lecture Notes in Statistics 31, Springer Verlag, New
York, 1985

Serfling, R.J., Approximation Theorems of Mathematical Statistics,
Wiley, New York, 1980

Spallek, K., Kurven und Karten, Bibliographisches Institut,
Zürich, 1980

Strasser (1), H., Mathematical Theory of Statistics, de Gruyter
Studies in Mathematics 7, de Gruyter, Berlin, 1985

Strasser (2), H., Einführung in die lokale asymptotische Theorie
der Statistik, Bayreuther Mathematische Schriften 19, Math.
Institut der Universität Bayreuth, 1985

Witting, H., Mathematische Statistik I, Teubner, Stuttgart, 1985

Wolfowitz, J., Asymptotic Efficiency of the Maximum Likelihood
Estimator, Theory Prob. and Applic. 10, 2, 1965

Lebenslauf
des
Bernd Kuhle
Am Waldschlösschen 32
4630 Bochum

27.02.1959 geboren

Juni 1977 Abitur, Theodor-Körner-Schule, 4630 Bochum

Sep. 1977 Einschreibung bei der Ruhr-Universität Bochum
für die Fächer Mathematik mit Nebenfach
Wirtschaftswissenschaft und Philosophie

März 1979 Vordiplom in Mathematik, Ruhr-Universität
Bochum

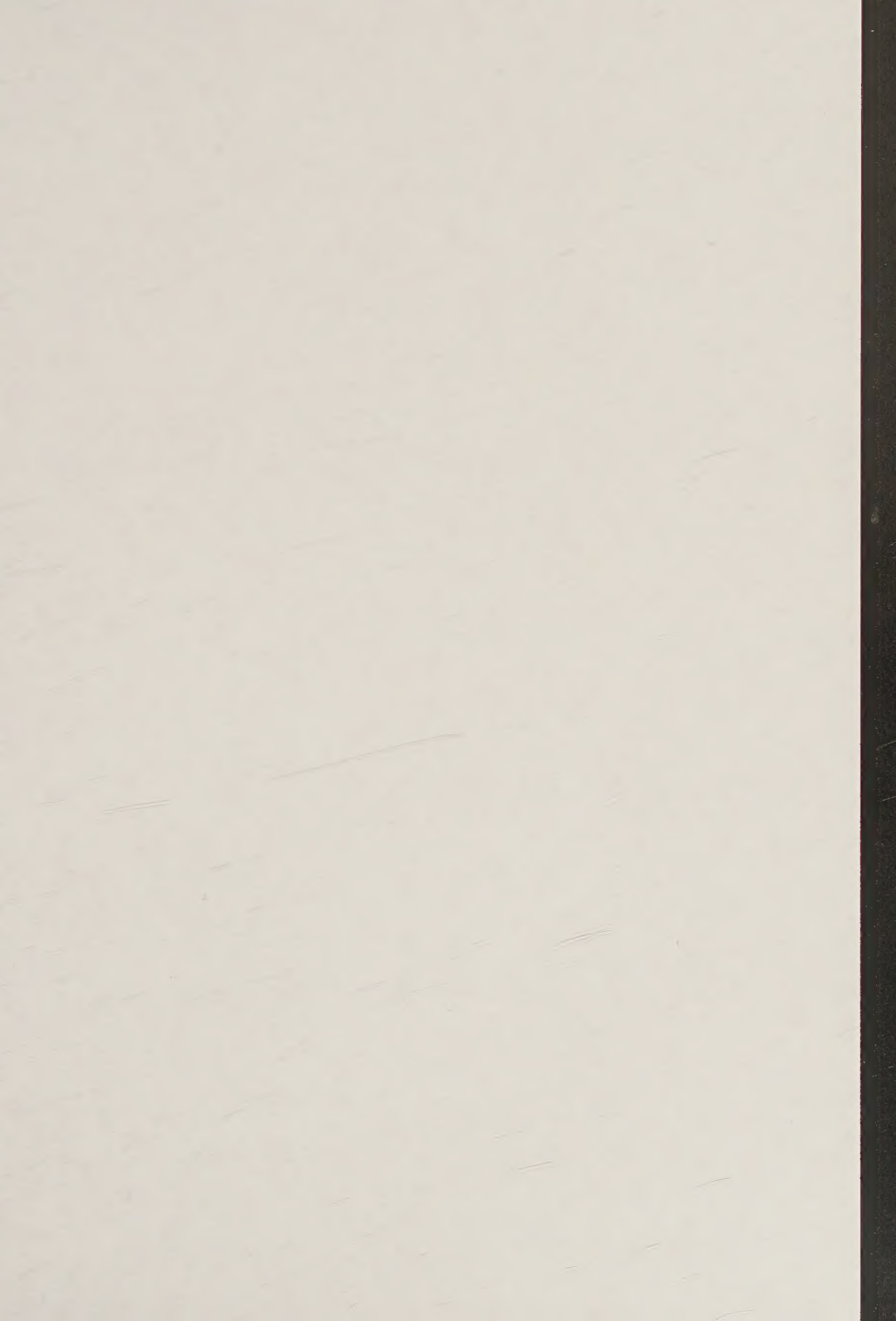
Juli 1979 bis
Juni 1983 Studentische Hilfskraft am Rechenzentrum der
Ruhr-Universität Bochum

Jan. 1984 Diplom in Mathematik, Ruhr-Universität Bochum

Jan. 1984 bis
Juni 1984 selbständiger Diplom-Mathematiker, Forschungs-
projekte für ein Wirtschaftsforschungsinstitut

Juli 1984 bis
Feb. 1986 Zivildienst

Feb. 1986 wissenschaftliche Hilfskraft am Lehrstuhl für
mathematische Statistik an der Ruhr-
Universität Bochum, Prof. Dr. V. Baumann.



D. GRÄBNER

Dissertations- u. Offsetdruckerei

Gotenstraße 3

8602 Altendorf bei Bamberg

Telefon 09545 / 8791

